



US006118472A

United States Patent [19]

Dureau et al.

[11] **Patent Number:** 6,118,472[45] **Date of Patent:** *Sep. 12, 2000

[54] **METHOD AND APPARATUS FOR SEAMLESS CONNECTIVITY OF WIDE-BAND NETWORKS AND NARROW-BAND NETWORKS**

[75] **Inventors:** Vincent Louis Dureau, Menlo Park, Calif.; Regis St. Girons, Fourqueux, France

[73] **Assignees:** Sun Microsystems, Inc., Palo Alto; OpenTV, Inc., Mountain View, both of Calif.

[*] **Notice:** This patent issued on a continued prosecution application filed under 37 CFR 1.53(d), and is subject to the twenty year patent term provisions of 35 U.S.C. 154(a)(2).

[21] **Appl. No.:** 08/658,498

[22] **Filed:** Jun. 5, 1996

[51] **Int. Cl.⁷** H04N 7/10; H04N 7/14

[52] **U.S. Cl.** 348/6; 455/3.1; 455/3.2; 455/5.1; 455/6.1; 348/7; 348/12; 348/13; 379/90.01

[58] **Field of Search** 348/6, 7, 8, 10, 348/12, 13; 455/3.1, 4.1, 4.2, 5.1, 6.1, 6.2, 3.2; 379/90.01

[56] **References Cited**

U.S. PATENT DOCUMENTS

5,448,568 9/1995 Delpuch et al. 348/14
5,481,542 1/1996 Logston 348/7

5,534,913 7/1996 Majeti et al. 348/7
5,572,442 11/1996 Schulhof et al. 348/6
5,586,121 12/1996 Moura et al. 348/12
5,608,447 3/1997 Farry et al. 348/7
5,612,730 3/1997 Lewis 348/13
5,629,978 5/1997 Blumhardt et al. 455/4.2
5,635,918 6/1997 Tett 340/825.52
5,640,453 6/1997 Schuchman et al. 348/12
5,694,163 12/1997 Harrison 348/13

Primary Examiner—John K. Peng

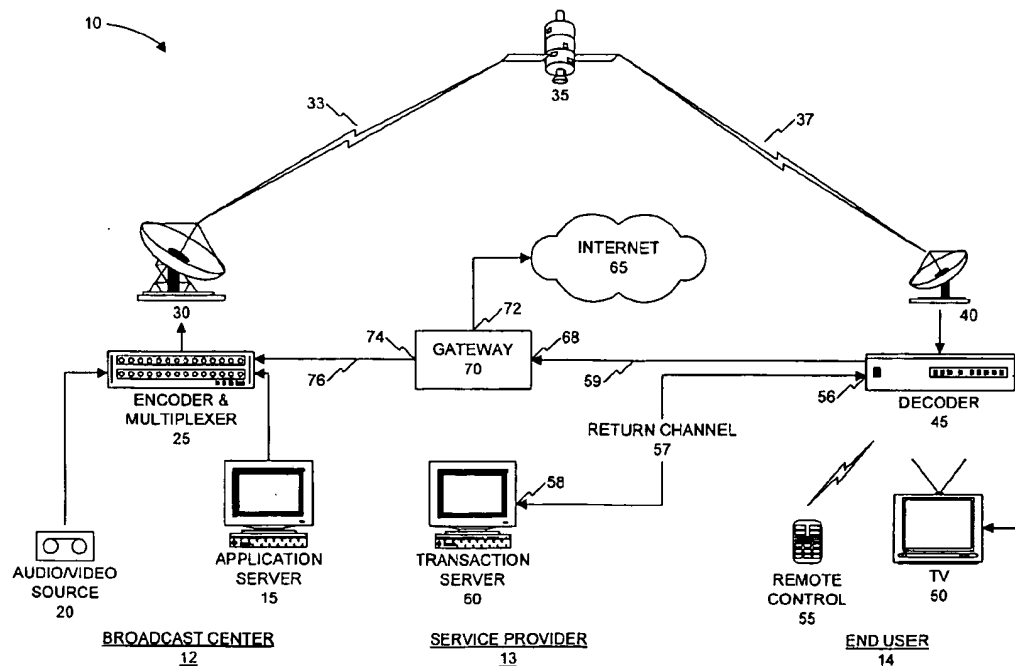
Assistant Examiner—Linus H. Lo

[57]

ABSTRACT

An apparatus for seamless connectivity between a narrow-band network like the Internet and an interactive TV wide-band network and methods of operating the same support on the fly translation and routing of data between the Internet and the interactive TV wide-band network. The apparatus for interfacing between a wide-band network and a narrow-band network comprises a decoder, having a decoder input, a first decoder port, and a second decoder port, which receives wide-band data from the wide-band network via the second decoder port and decodes the wide-band data to provide decoded data in response to decoder requests from the decoder input. A gateway, coupled to the decoder, the narrow-band network, and the wide-band network having a first gateway port to receive the decoder requests from the first decoder port, a second gateway port to interface with the narrow-band network, and a third gateway port to interface with the wide-band network, retrieves requested data from the narrow-band network in response to the decoder requests and transfers the requested data to the wide-band network for transfer to the decoder.

15 Claims, 3 Drawing Sheets



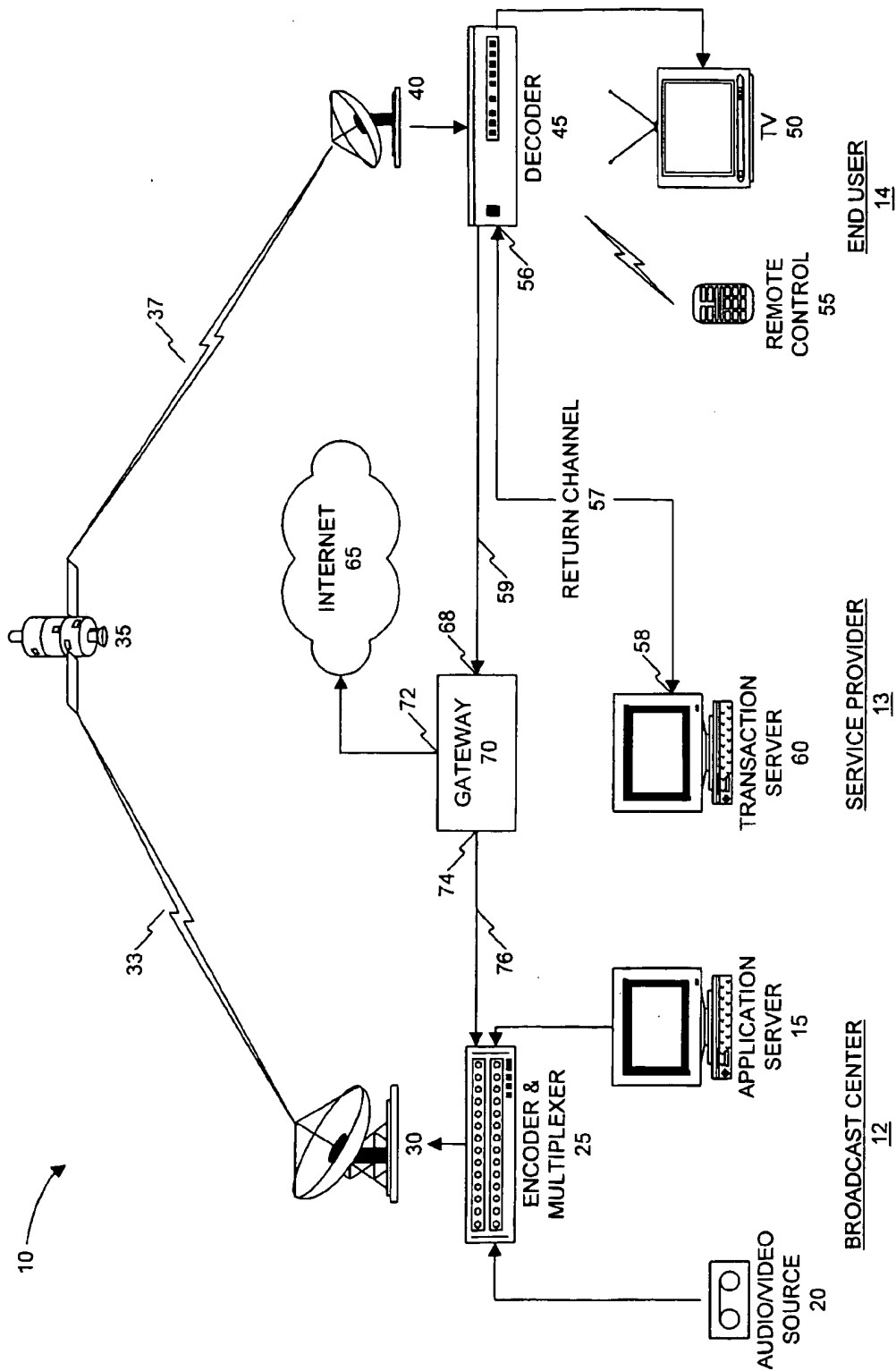


FIG. 1

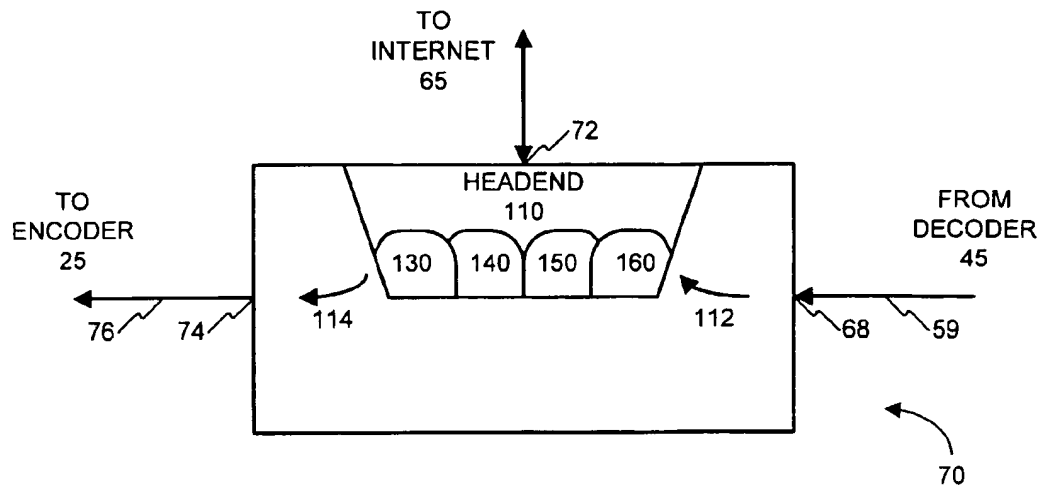


FIG. 2

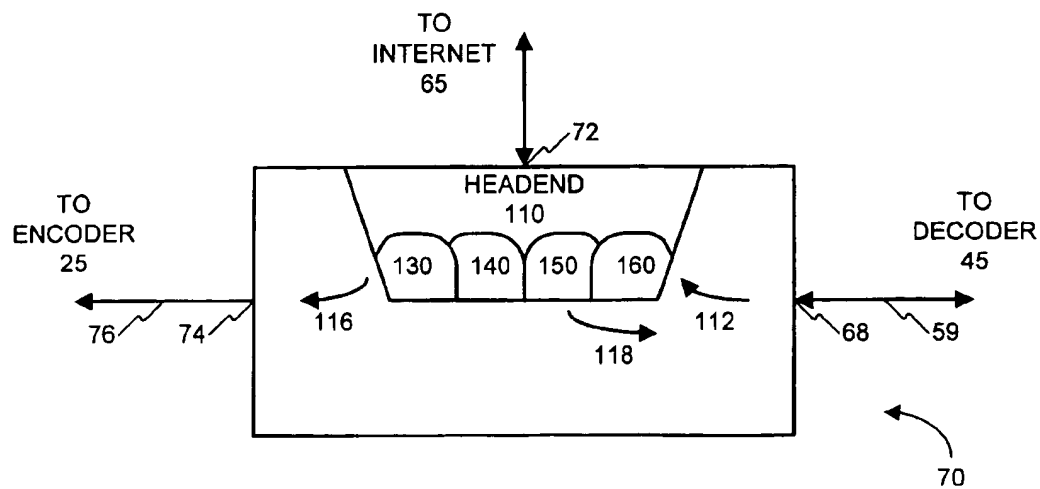


FIG. 3

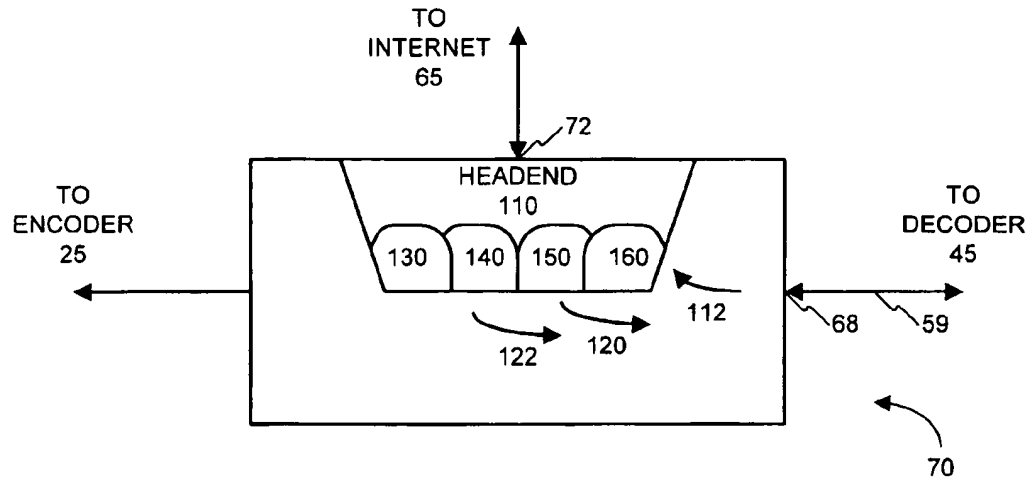


FIG. 4

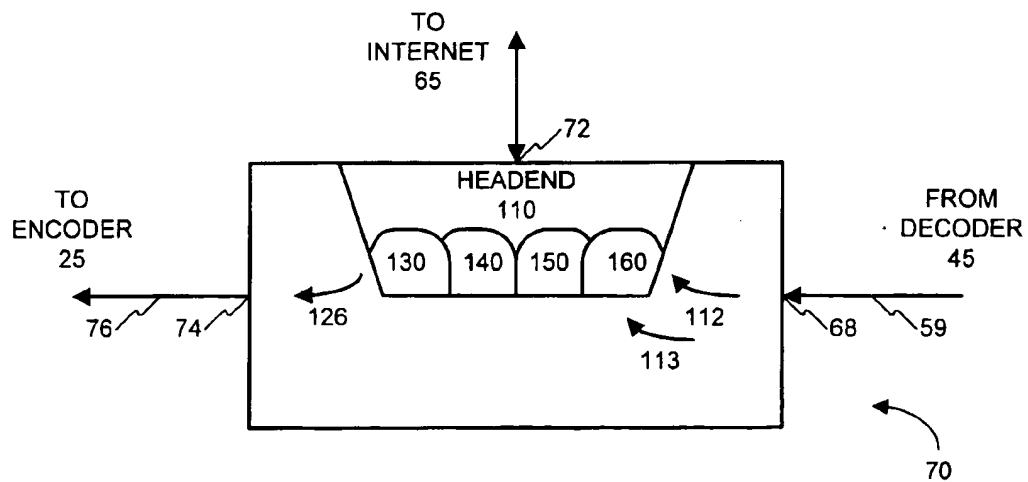


FIG. 5

1

METHOD AND APPARATUS FOR SEAMLESS CONNECTIVITY OF WIDE-BAND NETWORKS AND NARROW-BAND NETWORKS

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention generally relates to interfacing a wide-band network with a narrow-band network and more particularly to transparently interface a unidirectional wide-band broadcast network and the Internet.

2. Description of the Related Arts

Currently, there are two dominant digital infrastructures that are widely accepted and in public use. The first is the Internet structure also known as the world wide web based on narrow-band networks, and the second is Digital Pay television (TV) networks such as DirectTV™ based on wide-band networks.

Being the two dominant digital infrastructures, a single super hybrid infrastructure would provide immense amount of information for its users. Users having access to the single super hybrid infrastructure would have limitless options available. However, no provisions have been made to converge the two infrastructures into a single super hybrid infrastructure. Moreover, considering the properties of each of these networks provides an understanding as to why the two infrastructure are not likely to converge into a single super hybrid infrastructure.

The Internet infrastructure relies on a backbone of limited bandwidth in view of the number of users and services that the Internet infrastructure supports. Users are typically limited to 28.8 kb/s (kilo bits/sec) accessing through telco lines. A fraction of the users are able to upgrade to cable modems capable of Mb/s (mega bits/sec) transfer rates. The terminals used to access the Internet possess high processing power and large amounts of storage. These terminals are commonly referred to as desktop computers. The terminal displays of these desktop computers also possess the ability to produce high quality pictures. In an effort to take advantage of the high processing power of the terminals and reduce bandwidth consumption of the Internet, programming for the Internet relies on large amounts of computer caching available at the terminals and are relying increasingly on distributed processing which downloads a portion of the processing into the terminal to complete the processing for accessing the various world wide web sites.

On the other hand, the wide-band network based Digital Pay TV networks relies on a wide-band broadcast mono-directional network combined with a point to point low bit rate (2400 bits/sec) bi-directional network. The terminals which receive the wide-band broadcast data possess low processing power and little to no storage medium. The displays coupled to the terminals are low quality interlaced displays such as a typical National Television Standards Committee (NTSC) TV found commonly in most households. Thus, the terminals available to the wide-band networks possess low processing power with virtually no storage medium for data and provide low quality displays compared to the Internet terminals.

Further advances based on the Digital Pay TV networks include interactive TV systems. In order to make interactive TV less costly and therefore more attractive to consumer acceptance, it is desirable to keep memory requirements in the receiver to a minimum. Thus, as development for interactive TV systems continues, the trend is to continue building terminals with low processing power and low storage requirements.

2

Contrary to the trend of maintaining low processing power and minimizing storage requirements, new Internet protocols that are being developed rely more on the processing power of the latest generation computers. Thus, as the Internet technology based on the narrow-band network develops and the interactive TV technology based on the wide-band network technology develops, the Internet technology and the interactive TV technology continues to alienate each other and move apart. Consequently, consumers seeking Internet access that subscribe to interactive TV are forced to acquire the latest generation computer. The cost of having both a subscription to interactive TV and the latest generation computer for Internet access can be cost prohibitive for the consumer. In order for the providers of interactive TV to supply a complete service, a transparent consumer interface between the wide-band network of interactive TV and the narrow-band network for the Internet is needed.

Therefore, it is desirable to provide Internet connectivity to low end terminals operating in an interactive TV wide-band network, and a method of operating the same that support on the fly translation and routing of data between the Internet and the interactive TV wide-band network for transparent access to the Internet.

SUMMARY OF THE INVENTION

The present invention provides an apparatus for seamless connectivity between the Internet and an interactive TV wide-band network and methods of operating the same which support transparent on the fly translation and routing of data between the Internet and the interactive TV wide-band network. The novel seamless connectivity between the networks is based on a gateway that provides the translation and routing of data. Thus, according to one aspect of the invention, a seamless connection for interfacing between a wide-band network and a narrow-band network comprises a decoder, having a decoder input, a first decoder port, and a second decoder port, which receives wide-band data from the wide-band network via the second decoder port and decodes the wide-band data to provide decoded data in response to decoder requests from the decoder input. A gateway, coupled to the decoder, the narrow-band network, and the wide-band network having a first gateway port to receive the decoder requests from the first decoder port, a second gateway port to interface with the narrow-band network, and a third gateway port to interface with the wide-band network, retrieves requested data from the narrow-band network in response to the decoder requests and transfers the requested data to the wide-band network for transfer to the decoder. The decoder receives the requested Internet data from the wide-band network.

According to one aspect of the invention, the seamless connection further comprises a display device having a display input which displays the requested data and wherein the decoder includes a third decoder port coupled to the display input to provide decoded requested data to the display device. Thus, the requested Internet data is displayed on the display device.

~~According to another aspect of the invention, the first decoder port includes a bi-directional port, and the first gateway port includes a bi-directional port. The gateway includes circuitry which parses the requested data to provide narrow-band data and wide-band data and transfers the narrow-band data for output to the first decoder port. The decoder receives the narrow-band data and decodes the narrow-band data for output to the display device. The gateway transfers the wide-band data to the wide-band~~

3

network. The decoder receives the wide-band data and recombines the wide-band data with the narrow band data for output to the display device. ~~Thus, criteria may be established to efficiently determine the dynamic routing of the requested data between the wide-band network and the low bit rate network to the decoder.~~ Criteria include the type of the requested Internet data, the availability of bandwidth, size of the Internet data, and added costs associated with transmitting using the wide-band network.

An apparatus and method for seamless connectivity between the Internet and an interactive TV wide-band network are provided. The gateway supports a high performance computer for executing the native protocols of the Internet. The gateway parses the Internet data which enables the low processing power decoders to process the Internet data for display. Thus, lower cost decoders with low processing power can function to provide Internet access.

Other aspects and advantages of the present invention can be seen upon review of the figures, the detailed description, and the claims which follow.

BRIEF DESCRIPTION OF THE DRAWING

FIG. 1 illustrates a system level block diagram for digital transmissions in a direct broadcast satellite system for a wide-band network.

FIG. 2 illustrates a block diagram of an embodiment of a gateway transferring Internet data via the wide-band network according to the present invention.

FIG. 3 illustrates a block diagram of another embodiment of the gateway transferring Internet data via the wide-band network and the communication channel according to the present invention.

FIG. 4 illustrates a block diagram of another embodiment of the gateway transferring Internet data via the communication channel according to the present invention.

FIG. 5 illustrates a block diagram of another embodiment of the gateway transferring a plurality of Internet data for repeated broadcasts via the wide-band network according to the present invention.

DETAILED DESCRIPTION OF THE INVENTION

The invention will be described with respect to the Figs. in which FIG. 1 generally shows a digital transmission system, as for example a direct broadcast satellite system. It is presumed that a single satellite transponder accommodates a plurality of respective TV programs in time division multiplexed format.

Referring to FIG. 1, the direct broadcast satellite system 10 provides a wide-band network that includes a broadcast center 12, service provider 13, and an end user 14. The broadcast center 12 includes an application server 15, an audio/video source 20, encoder and multiplexer 25, and satellite transmitter 30. The application server 15 controls execution of interactive TV applications which are loaded into the logic circuits of the application server 15 to perform a series of specifically identified operations dictated by the interactive TV applications.

The interactive TV applications include associated audio and video information sources 20. The application server 15 synchronizes the interactive TV applications and the associated audio and video information sources 20 into transport packets that provide inputs to the encoder and multiplexer 25. The encoder and multiplexer 25 receives the transport packets and encodes the transport packets for transmission.

4

Satellite transmitter 30 time-multiplexes the transport packets and transmits the transport packets as upload signal 33 to satellite 35.

The broadcast center 12 is described in detail in issued U.S. Pat. No. 5,448,568 entitled "System of Transmitting an Interactive TV Signal" assigned to Thomson Consumer Electronics, Inc. issued Sep. 5, 1995. The U.S. Pat. No. 5,448,568 is herein incorporated by reference in its entirety.

Satellite 35 receives the upload signal 33 and transmits download signal 37 to end user 14. The end user 14 includes satellite dish 40, decoder 45, TV 50, remote control 55, return channel 57, and communication channel 59. Satellite dish 40 receives the download signal 37 and provides an output to decoder 45. The decoder 45 includes a software operating system loaded into the logic circuits of the decoder 45 that performs a series of steps to control the operations of the decoder 45. The decoder 45 receives the download signal 37 from satellite dish 40 and decodes the transmitted interactive TV application and its associated audio and video information 20. The decoder 45 executes the interactive TV application and provides audio and video outputs to TV 50.

Remote control 55 provides inputs to the decoder 45 to select execution of other interactive TV applications for output to TV 50. The decoder 45 includes an input/output port 56 that couples to the return channel 57 for communication to transaction server 60 or the communication channel 59 for communication with gateway 70.

The service provider 13 provides local interaction with the end user 14 and includes the transaction server 60 and the gateway 70. The transaction server 60 includes an input/output port 58 which couples to the return channel 57. The transaction server 60 provides monitoring of transactions performed by the end user 14 and updating of the software operating system for the decoder 45 via the return channel 57. The gateway 70 includes a port 68 which receives request for Internet data from decoder 45 via the communication channel 59. The return channel 57 and the communication channel 59 may be telephone lines or cable lines and support a low bit rate link.

The gateway 70 includes a port 72 which provides access to Internet 65. The Internet 65 is a narrow-band network commonly known as the world wide web. The gateway 70 retrieves Internet data from the Internet 65 and communicates the Internet data to port 74 of the gateway 70. High speed line 76 transfers the Internet data to the encoder and multiplexer 25 which encodes the Internet data for broadcast to the wide-band network. Decoder 45 receives the encoded Internet data and decodes the Internet data for display on TV 50.

FIG. 2 illustrates a block diagram of an embodiment of the gateway 70 according to the present invention. The gateway 70 includes a headend 110 that includes circuitry programmed to execute native protocols of the Internet 65. The headend 110 includes an MPEG (Motion Picture Experts Group) encoder 130, renderer 140, cache 150, and parser 160. Port 68 of the gateway 70 receives a data request 112 from decoder 45. Communication channel 59 transfers the data request 112 from the decoder 45 to the port 68. Communication channel 59 is a low bit rate link utilizing telephone lines as the medium for data transfers from the decoder 45 and the gateway 70. The gateway 70 receives the data request 112 for Internet data and forwards the data request 112 to the headend 110.

The headend 110 executes the native protocols for the data request 112 and receives Internet data 114 from the Internet 65. The cache 150 manages latency between the

Internet web servers and the headend 110 during retrieval of Internet data 114. The headend 110 transfers the Internet data 114 to port 74 where the high speed line 76 transfers the Internet data 114 to the encoder and multiplexer 25 for encoding. The Internet data 114 includes spatial and temporal correlation encoded data such as rendered MPEG encoded audio and video data. The MPEG encoder 130 and renderer 140 provides the rendered MPEG graphical data. The Internet data 114 is encoded for broadcast by the satellite 30 to the wide-band network. The decoder 45 receives the encoded Internet data and decodes the Internet data for display by TV 50.

Given that the MPEG encoder 130 and render 140 process the Internet data 114, the decoder 45 does not require high processing power to render the graphical data from the Internet data 114. The headend 110 provides a simple and high quality graphical interface to the Internet. Flexibility in MPEG encoding and rendering is also achieved. Moreover, the nature of MPEG encoding takes advantage of changes in relation to static portions of a page. Once the static portions of a page is transferred, later transfers provide data that represent changes to the static portions of the page. Thus, various tradeoffs are considered for the transfer of the MPEG data to the decoder such as fixed quality encoding with variable latency or variable quality with fixed latency. Other tradeoffs include fixed bandwidth verses variable bandwidth allocation of the wide-band link. The tradeoffs are taken into account for minimizing the processing of Internet data for the decoder 45.

FIG. 3 illustrates a block diagram of another embodiment of the gateway 70. The gateway 70 includes a headend 110 that includes circuitry programmed to execute native protocols of the Internet 65. Port 68 of the gateway 70 receives a data request 112 from decoder 45. Communication channel 59 transfers the data request 112 from the decoder 45 to the port 68. Communication channel 59 is a low bit rate link utilizing telephone lines as the medium for data transfers between the decoder 45 and the gateway 70. The gateway 70 receives the data request 112 for Internet data and forwards the data request 112 to the headend 110.

The headend 110 executes the native protocols for the data request 112 and receives Internet data 114 from the Internet 65. The parser 160 parses the Internet data 114 into wide-band data and narrow-band data such as graphical data and textual data. The headend 110 determines routing of the parsed Internet data between the wide-band links and the narrow-band links based on criteria such as cost, availability of bandwidth, size of data, and type of data. For example, small sized data may be transferred using the slower low bit rate link. Real time data on the other hand such as video or audio data are transferred using the faster wide-band network.

Referring to FIG. 3, the gateway 70 receives the wide-band Internet data 116 and transfers the wide-band Internet data 116 to port 74 where the high speed line 76 transfers the wide-band Internet data 116 to the encoder and multiplexer 25 for encoding. After encoding the wide-band Internet data 116, satellite 30 broadcasts the encoded Internet data to the wide-band network. The gateway 70, in response to the narrow-band data 118, transfers the narrow-band Internet data 118 to port 68 where the communication channel 59 transfers the narrow-band Internet data 118 to the decoder 45. The decoder 45 receives the encoded wide-band Internet data 116 from the wide-band network and the narrow-band Internet data on communication channel 59. After decoding the encoded wide-band Internet data 116, the decoder 45 includes circuitry which recombines the wide-band Internet

data 116 with the narrow-band Internet data 118 to provide display data for display by TV 50.

The decoder 45 includes circuitry that renders the graphical data for display by TV 50. Bandwidth consumption of the wide-band network is reduced by parsing the internet data to provide graphical data and textual data and using the narrow-band network to transfer the parsed internet data.

FIG. 4 illustrates another embodiment of the gateway 70 according to the present invention. As decoder 45 receives a request to access the Internet 65, the decoder tunes to a specific channel on the wide-band network and downloads an interactive TV application associated with Internet access. The interactive TV application includes an Internet web browser similar to the Netscape Navigator™ developed by Netscape Communications Corporation and instruction codes that programs circuitry within the decoder 45 to establish the communication channel 59 to contact the gateway 70. Thus, the decoder 45 includes additional circuitry for increased processing power that enables execution of a world wide web browser at the decoder 45.

As the decoder 45 receives inputs for Internet requests 112, the Internet requests 112 are transferred to port 68 of the gateway 70 via the established communication channel 59. The gateway 70 transfers the Internet requests 112 to the headend 110. The headend 110 includes circuitry programmed to retrieve the requested Internet data. Once the requested Internet data is retrieved, the headend 110 parses the Internet data to provide text 120 and control devices 122 for the particular web page. The control devices 122 include functional buttons and scrolling functions for the web page.

The gateway 70 transfers the text 120 and the control devices 122 to the decoder 45 via communication channel 59. The text 120 and the control 25 devices 122 provide inputs to the Internet web browser which enables the decoder 45 to reconstruct the requested Internet data. As the decoder 45 request additional web pages, the headend 110 transfers additional text 120 and control devices 122 associated with the additional web pages via the communication channel 59. Thus, as the decoder possesses more processing power, the amount of data and the rate of transfer for the data to support Internet browsing is reduced.

FIG. 5 illustrates another embodiment of the gateway 70 according to the present invention. During instances when requests for access to the Internet 65 are abundant, certain web pages on the Internet 65 are more popular and are repeatedly requested. The headend 110 includes the cache 150 and circuitry programmed to statistically monitor and store repeated retrievals of the more popular web pages from the Internet 65. Data from the statistical monitoring establish controls for rebroadcasting the more popular web pages to the wide-band network. Rebroadcasting the more popular web pages enables faster responses to decoders 45 that request those web pages.

For example, Internet requests 112 and 113 request similar Internet data and are received at port 68 via Communication channel 59 from a plurality of decoders 45. The gateway 70 forwards the Internet requests 112 and 113 to the headend 110. The headend 110 retrieves the requested Internet data 126 and determines that the Internet data is frequently requested and notifies the headend 110. As the headend 110 transfers the requested Internet data 126 to port 74 for transfer to the encoder 25 via the high speed line 76, the encoder 25 is also notified that the requested Internet data 126 is to be rebroadcasted. The encoder 25 encodes the requested Internet data 126 for broadcast by the satellite 30 and notifies the satellite to repeatedly broadcast the encoded

7

internet data. The satellite 30, in response to the notification, rebroadcasts the requested Internet data 125 received from cache 150 at regular intervals. Increased decoder response is achieved by having the popular web pages stored in cache 150 and regularly broadcasted rather than having each individual decoder submit requests to the gateway 70 to retrieve the same popular web pages.

In a further embodiment of the invention, the decoder 45 includes circuitry programmed to select particular web pages that have been requested from the repeated broadcasts of the popular web pages. The other web pages received from the wide-band network are filtered and ignored by the decoder 45. Thus, the decoder 45 is programmed to select from the popular web pages broadcasted onto the wide-band network those web pages that the decoder 45 had requested. Rebroadcasting the popular web pages enables the gateway 70 to operate at higher efficiency by avoiding repeated accesses to Internet 65. Furthermore, latency associated with accessing the Internet 65 is reduced by having the popular web pages readily download able from the wide-band network. Depending on the traffic for accessing particular web pages, latency associated with accessing the particular web pages can be of long duration. Thus, repeated broadcasts of popular web pages from the Internet 65 increases response time to the decoder 45 when accessing the Internet.

Accordingly, an apparatus for seamless connectivity of wide-band networks and narrow-band networks has been provided. The gateway provides an interface between the wide-band network and the narrow-band network. Accessing the Internet which resides on the narrow-band network from the wide-band network base is transparent to the users of the decoder on the wide-band network. Depending on the type of Internet data and the sophistication of the circuitry of the decoder, the Internet data is transferred over the wide-band network or the low bit rate communication channel of the decoder or a combination of both. Although the embodiments for seamless connectivity of wide-band networks and narrow-band networks have been disclosed with reference to an interactive TV system operating in a satellite transceiving network and the Internet, variations of the seamless connectivity interface according to the present invention are applicable in other network applications.

The foregoing description of a preferred embodiment of the invention has been presented for purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise forms disclosed. Obviously, many modifications and variations will be apparent to the practitioners skilled in the art. It is intended that the scope of the invention be defined by the following claims and their equivalents.

What is claimed is:

1. A seamless connection for interfacing between a wide-band network and a narrow-band network comprising:

- a decoder, having a decoder input, a first decoder port, and a second decoder port, which receives wide-band data from the wide-band network via the second decoder port and decodes the wide-band data to provide decoded data in response to decoder requests from the decoder input;
- a gateway, coupled to the the coder, the narrow-band network, and the wide-band network having a first gateway port to receive the decoder requests from the first decoder port, a second gateway port to interface with the narrow-band network, and a third gateway port to interface with the wide-band network, which retrieves requested data from the narrow-band network

8

in response to the decoder requests and transfers the requested data to the wide-band network for transfer to the decoder; and

- a display device having a display input which displays the requested data and wherein the decoder includes a third decoder port coupled to the display input to provide decoded requested data to the display device;

wherein:

- the first decoder port includes a bi-directional port;
- the first gateway port includes a bi-directional port;
- the gateway includes circuitry which parses the requested data to provide narrow-band data and wide-band data and transfers the narrow-band data to the first gateway port for output; and
- the decoder receives the narrow-band data via the first decoder port.

2. The seamless connection according to claim 1, wherein:

- the gateway transfers the wide-band data to the wide-band network; and
- the decoder receives the wide-band data and recombines the wide-band data with the narrow band data for output to the display device.

3. A seamless connection for interfacing between a wide-band network and a narrow-band network comprising:

- a decoder, having a decoder input, a first decoder port, and a second decoder port, which receives wide-band data from the wide-band network via the second decoder port and decodes the wide-band data to provide decoded data in response to decoder requests from the decoder input;
- a gateway, coupled to the decoder, the narrow-band network, and the wide-band network having a first gateway port to receive the decoder requests from the first decoder port, a second gateway port to interface with the narrow-band network, and a third gateway port to interface with the wide-band network, which retrieves requested data from the narrow-band network in response to the decoder requests and transfers the requested data to the wide-band network for transfer to the decoder; and
- a display device having a display input which displays the requested data and wherein the decoder includes a third decoder port coupled to the display input to provide decoded requested data to the display device;

wherein the decoder includes circuitry programmable to execute specified functions receives a browser application from the wide-band network which programs the circuitry in the decoder to receive parsed Internet data from the narrow-band network for output to the display device.

4. The seamless connection according to claim 3, wherein the gateway receives Internet data pertaining to a particular web page and parses the web page to provide the parsed Internet data for transfer to the decoder.

5. The seamless connection according to claim 4, wherein the parsed Internet data includes text and control devices to reconstruct the particular web page.

6. A connection system for a wide-band broadcast network for accessing a narrow-band network to provide Internet access comprising:

- a wide-band receiver, coupled to the wide-band network, which receives broadcasts of encoded wide-band broadcast data from the wide-band network;
- a decoder, coupled to the wide-band broadcast network having a first decoder input, a second decoder input, a

first decoder port, and a second decoder port, responsive to decoder control signals applied to the first decoder input provides decoder requests to the first decoder port and decodes wide-band data received from the second decoder input to provide display data to the second decoder port;

a gateway, coupled to the decoder and the narrow-band network having a first gateway port to receive the decoder requests, a second gateway port to transceive Internet data, and a third gateway port, which retrieves requested Internet data from the narrow-band network in response to the decoder requests and transfers the requested Internet data to the third gateway port;

an encoder, coupled to the third gateway port having an encoder input to receive the requested Internet data and an encoder output, which encodes the requested Internet data and provides the encoded wide-band broadcast data to the encoder output;

a wide-band transmitter, coupled to the wide-band network and the encoder output, which broadcasts the encoded wide-band broadcast data to the wide-band network; and

a display coupled to the second decoder port which receives the display data and provides an visual image in response to the display data;

wherein:

the gateway includes circuitry which parses the requested Internet data to provide narrow-band data and wide-band data and transfers the narrow-band data for output to the first decoder port and the wide-band data to the encoder input; and

the decoder receives the narrow-band data and the encoded wide-band broadcast data and includes circuitry which recombines the narrow-band data and the encoded wide-band broadcast data to provide display data to the second decoder port.

7. A connection system for a wide-band broadcast network for accessing a narrow-band network to provide Internet access comprising:

a wide-band receiver, coupled to the wide-band network, which receives broadcasts of encoded wide-band broadcast data from the wide-band network;

a decoder, coupled to the wide-band broadcast network having a first decoder input, a second decoder input, a first decoder port, and a second decoder port, responsive to decoder control signals applied to the first decoder input provides decoder requests to the first decoder port and decodes wide-band data received from the second decoder input to provide display data to the second decoder port;

a gateway, coupled to the decoder and the narrow-band network having a first gateway port to receive the decoder requests, a second gateway port to transceive Internet data, and a third gateway port, which retrieves requested Internet data from the narrow-band network in response to the decoder requests and transfers the requested Internet data to the third gateway port;

an encoder, coupled to the third gateway port having an encoder input to receive the requested Internet data and an encoder output, which encodes the requested Internet data and provides the encoded wide-band broadcast data to the encoder output;

a wide-band transmitter, coupled to the wide-band network and the encoder output, which broadcasts the encoded wide-band broadcast data to the wide-band network; and

a display coupled to the second decoder port which receives the display data and provides an visual image in response to the display data;

wherein:

the decoder includes circuitry programmable to execute specified functions receives a browser application from the wide-band network which programs circuitry in the decoder to receive parsed Internet data from the narrow-band network for output to the display device.

8. The connection system according to claim 7, wherein the gateway includes circuitry which provides internet browsing to retrieve the requested Internet data from the narrow-band network and parse the requested Internet data to provide the parsed Internet data to the decoder.

9. The connection system according to claim 8, wherein the gateway requested Internet data includes a particular web page and the parsed Internet data includes text and control devices to reconstruct the particular web page.

10. A connection system for a wide-band broadcast network for accessing a narrow-band network to provide Internet access comprising:

a wide-band receiver, coupled to the wide-band network, which receives broadcasts of encoded wide-band broadcast data from the wide-band network;

a decoder, coupled to the wide-band broadcast network having a first decoder input, a second decoder input, a first decoder port, and a second decoder port, responsive to decoder control signals applied to the first decoder input provides decoder requests to the first decoder port and decodes wide-band data received from the second decoder input to provide display data to the second decoder port;

a gateway, coupled to the decoder and the narrow-band network having a first gateway port to receive the decoder requests, a second gateway port to transceive Internet data, and a third gateway port, which retrieves requested Internet data from the narrow-band network in response to the decoder requests and transfers the requested Internet data to the third gateway port;

an encoder, coupled to the third gateway port having an encoder input to receive the requested Internet data and an encoder output, which encodes the requested Internet data and provides the encoded wide-band broadcast data to the encoder output;

a wide-band transmitter, coupled to the wide-band network and the encoder output, which broadcasts the encoded wide-band broadcast data to the wide-band network; and

a display coupled to the second decoder port which receives the display data and provides an visual image in response to the display data;

wherein:

the gateway retrieves a plurality of web pages from the narrow-band network and transfers the plurality of web pages to the encoder;

the encoder encodes the plurality of web pages and provides encoded wide-band web page broadcast data to the wide-band transmitter; and

the wide-band transmitter transmits the wide-band web page broadcast data to the wide-band network.

11. The connection system according to claim 10, wherein the decoder includes circuitry programmed to select particular web pages from the wide-band web page broadcast data for decoding to provide display data for output.

12. A method of seamless connectivity on a hybrid network for interfacing between a wide-band network and a narrow-band network comprising the steps:

11

interfacing to the narrow-band network to retrieve narrow-band data in response to decoder requests; encoding the narrow-band data for transfer to the wide-band network; decoding wide-band data from the wide-band network to provide display data; and parsing the narrow-band data to provide low bit rate data and wide-band data.

13. A method of seamless connectivity on a hybrid network for interfacing between a wide-band network and a narrow-band network comprising the steps:

interfacing to the narrow-band network to retrieve narrow-band data in response to decoder requests; encoding the narrow-band data for transfer to the wide-band network; decoding wide-band data from the wide-band network to provide display data; parsing the narrow-band data to provide low bit rate data and wide-band data; transferring the low bit rate data to a low bit rate link; receiving the low bit rate data from the low bit rate link; and recombining the low bit rate data with the wide-band data to provide the display data.

14. A method of seamless connectivity on a hybrid network for interfacing between a wide-band network and a narrow-band network comprising the steps:

12

interfacing to the narrow-band network to retrieve narrow-band data in response to decoder requests; encoding the narrow-band data for transfer to the wide-band network; decoding wide-band data from the wide-band network to provide display data; parsing the narrow-band data to provide low bit rate data and wide-band data; downloading a client browser application from the wide-band network; parsing the narrow-band data to provide low bit rate data to the client browser application; and processing the low bit rate data to provide the display data.

15. The method of seamless connectivity according to claim 14, wherein:

the low bit rate data includes image rectangles representing optional wide-band data; transferring the optional wide-band data to the wide-band network; and recombining the optional wide-band data with the low bit rate data of the client browser application to display the optional wide-band data in response to a decoder request.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE
CERTIFICATE OF CORRECTION

PATENT NO : 6,118,472

DATED : Sep. 12, 2000

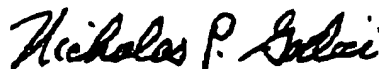
INVENTOR(S): Dureau, et al.

It is certified that an error appears in the above-identified patent and that said Letters Patent are hereby corrected as shown below:

In claim 1 (at column 7, line 61), replace "the the coder" with --the decoder--.

Signed and Sealed this
Twenty-fourth Day of April, 2001

Attest:



NICHOLAS P. GODICI

Attesting Officer

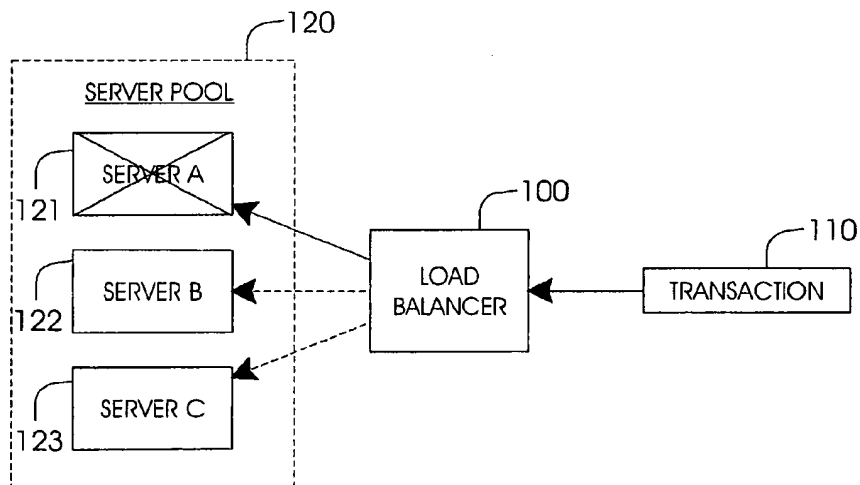
Acting Director of the United States Patent and Trademark Office



US 20020087694A1

(19) **United States**(12) **Patent Application Publication** (10) **Pub. No.: US 2002/0087694 A1**
Daoud et al. (43) **Pub. Date: Jul. 4, 2002**(54) **APPARATUS AND METHOD FOR
IDENTIFYING A REQUESTED LEVEL OF
SERVICE FOR A TRANSACTION**(76) Inventors: **Raja Daoud**, Santa Clara, CA (US);
Francisco J. Romero, Plano, TX (US)Correspondence Address:
HEWLETT-PACKARD COMPANY
Intellectual Property Administration
P.O. Box 272400
Fort Collins, CO 80527-2400 (US)(21) Appl. No.: **09/751,009**(22) Filed: **Dec. 29, 2000****Publication Classification**(51) Int. Cl.⁷ **G06F 15/173; G06F 15/16**(52) U.S. Cl. **709/226; 709/203**(57) **ABSTRACT**

An apparatus for identifying a requested level of service for a transaction wherein the transaction may be processed in accordance with the requested level of service. The invention is preferably embodied in computer readable program code stored in suitable storage media, and comprises, program code for selecting the requested level of service for the transaction, and program code for assigning the requested level of service to the transaction. The transaction is preferably a packetized signal comprising at least a data packet having a service tag associated therewith, wherein the service tag includes the requested level of service. The requested level of service can be any suitable factors or combination thereof, and can be assigned at any point on the network. The service tag is read from the transaction using suitable program code (e.g., at a load balancer), and based on the requested level of service, the transaction is directed to and processed by a network device that is best able to provide the requested level of service.



*Transactions
Routed/Routed
based on
level of service
paid for*

FIG. 1

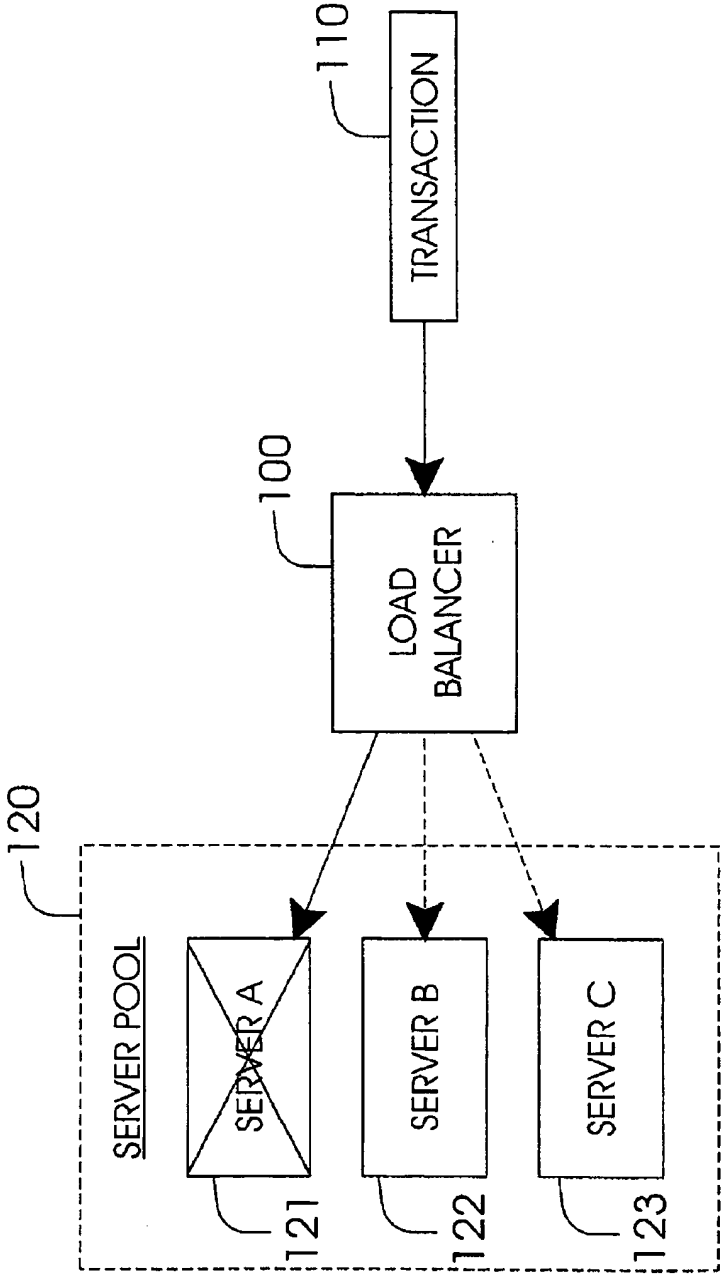


FIG. 2

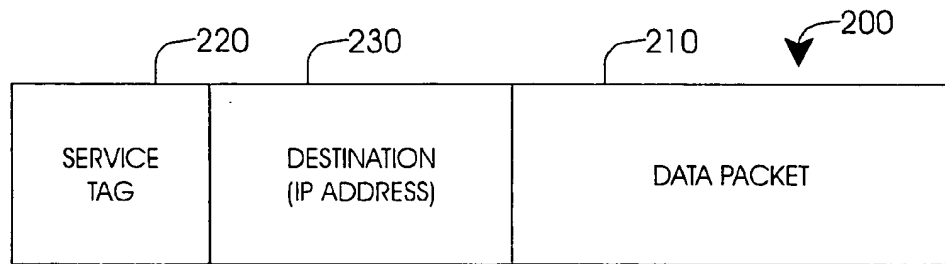


FIG. 4

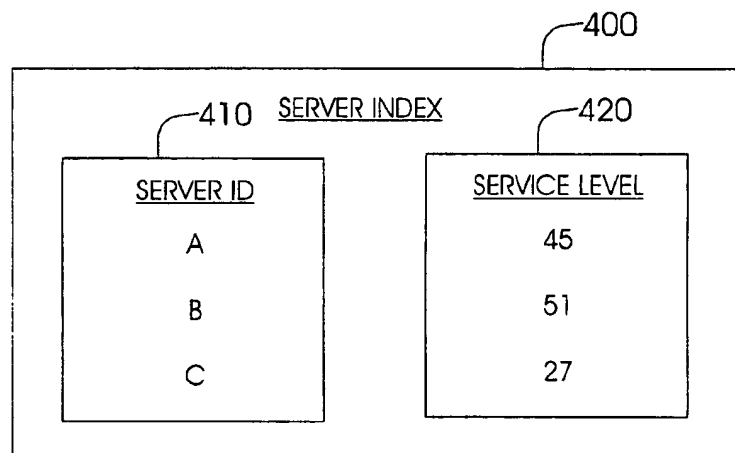


FIG. 6

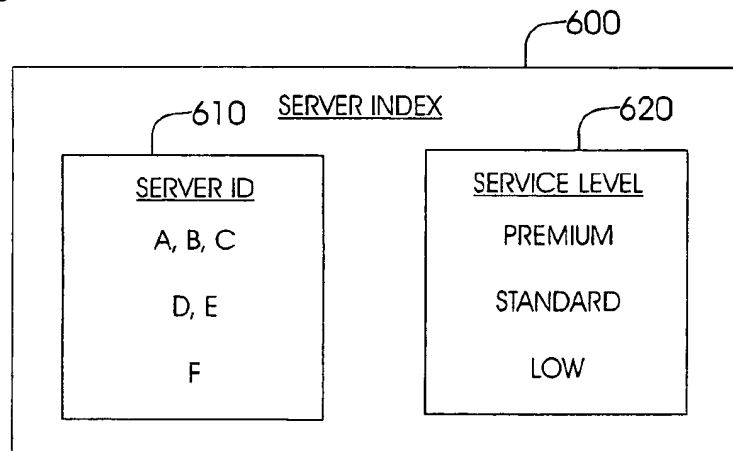


FIG. 3

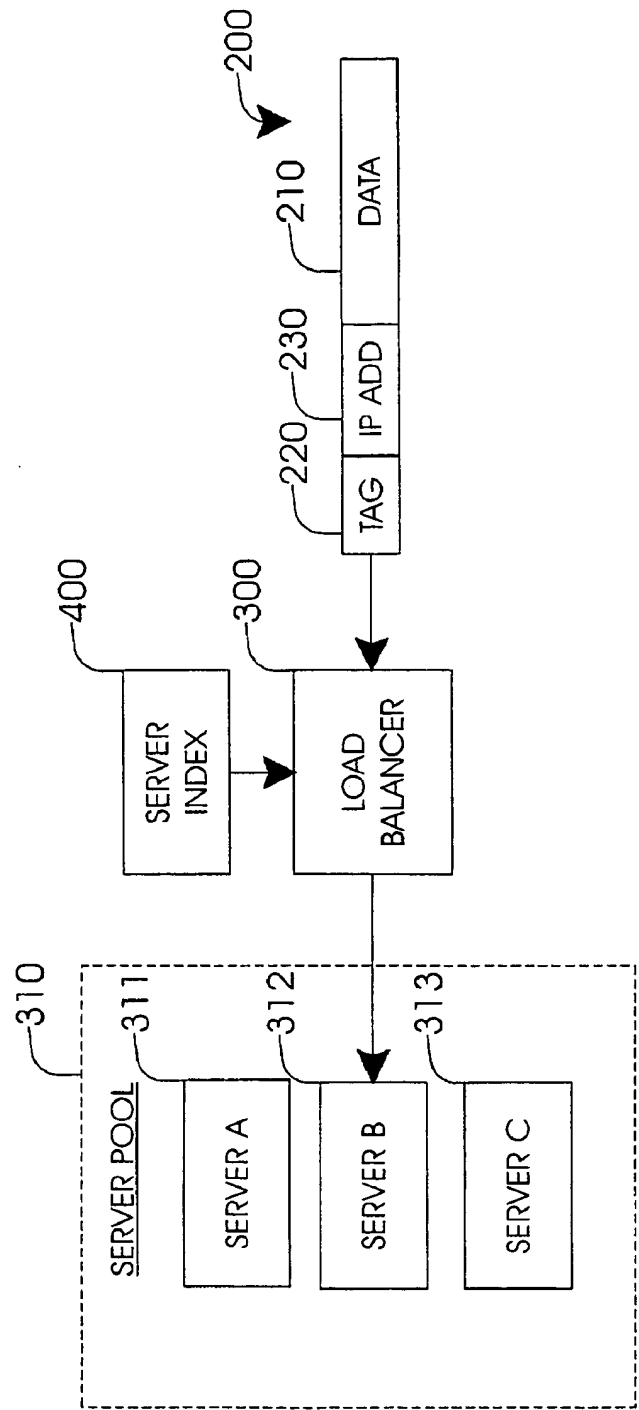


FIG. 5

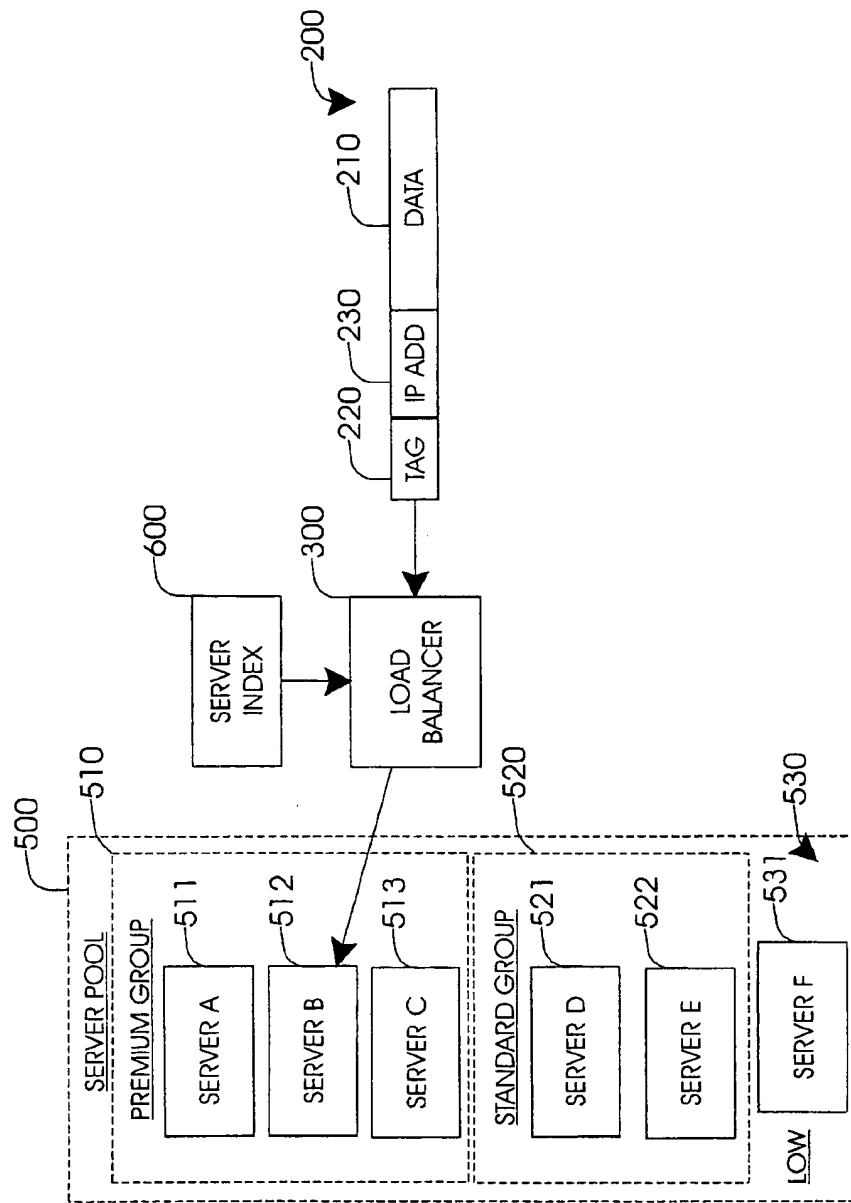


FIG. 7

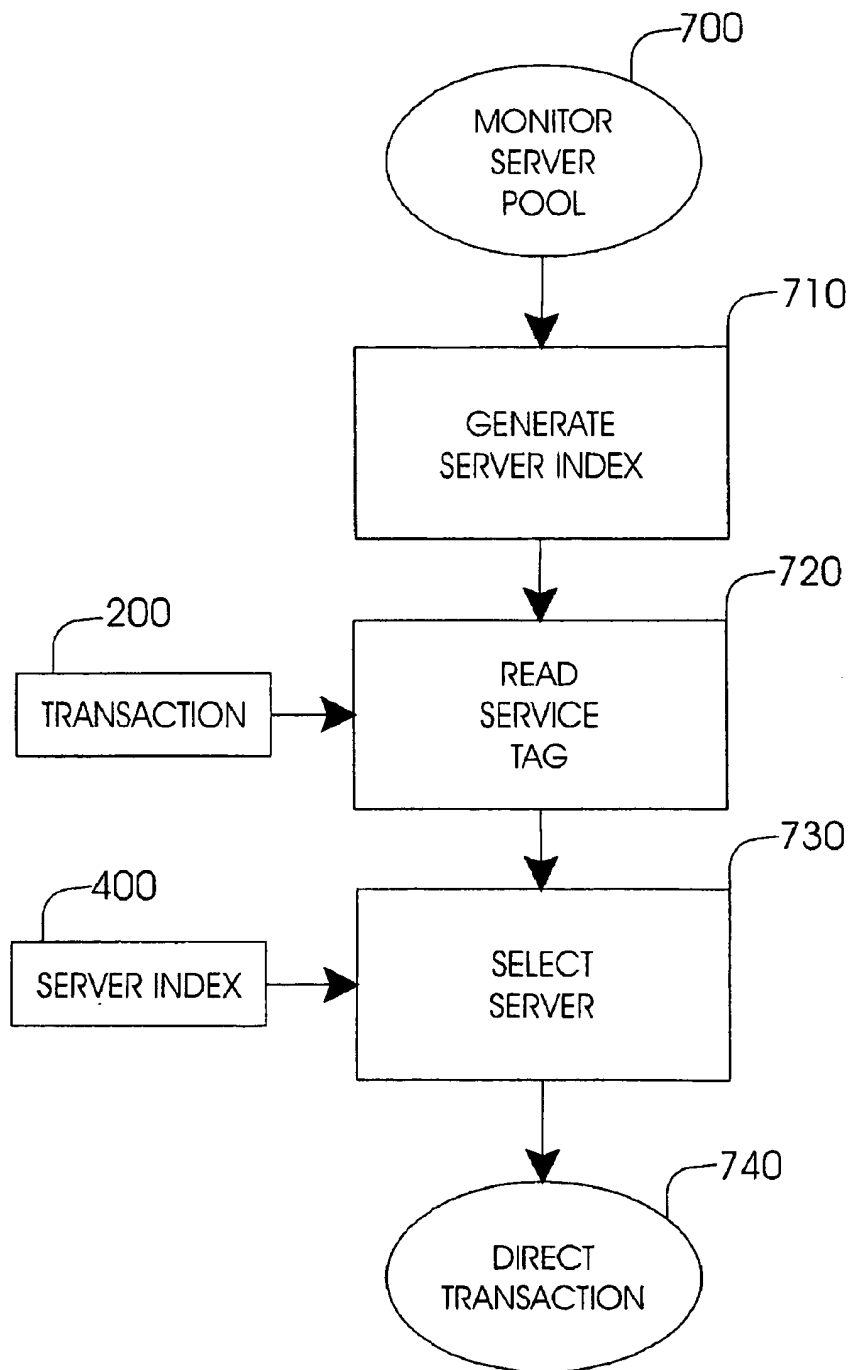


FIG. 8

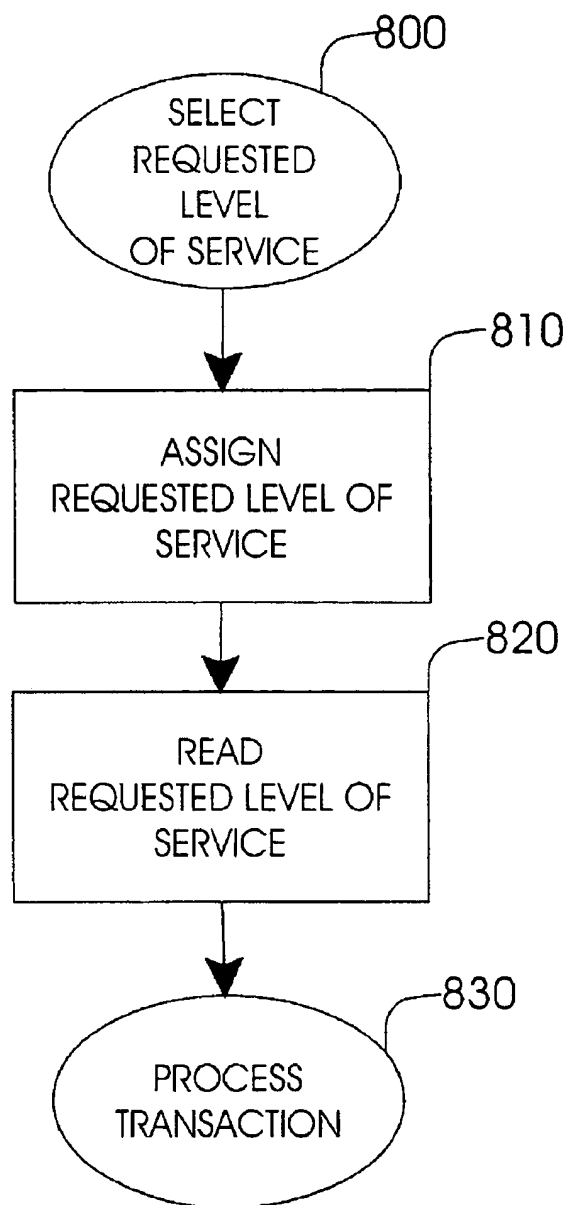
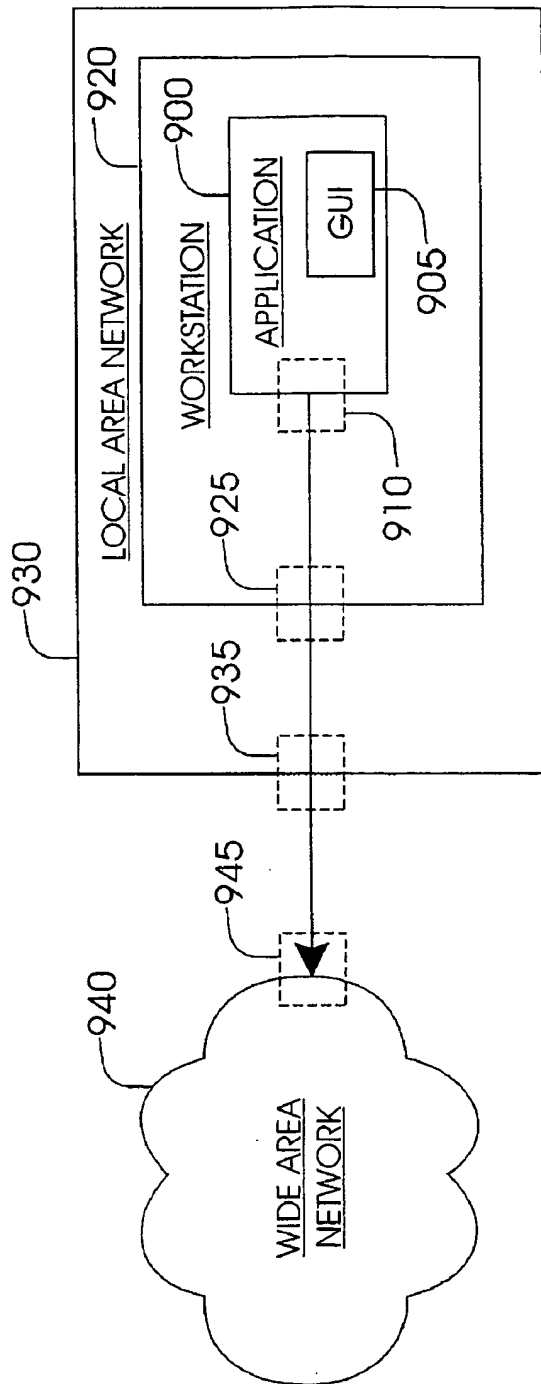


FIG. 9



APPARATUS AND METHOD FOR IDENTIFYING A REQUESTED LEVEL OF SERVICE FOR A TRANSACTION

RELATED APPLICATION

[0001] This patent application is related to co-owned patent application for APPARATUS AND METHOD FOR ROUTING A TRANSACTION BASED ON A REQUESTED LEVEL OF SERVICE, having the same filing date and identified by Hewlett Packard Docket No. HP 10002670-1.

FIELD OF THE INVENTION

[0002] The invention pertains to identifying a requested level of service for a transaction, wherein the transaction may be processed in accordance with the requested level of service.

BACKGROUND OF THE INVENTION

[0003] Server pools having multiple servers are often provided on networks, including the Internet, to handle large volumes of transactions (i.e., "requests to process data") thereon. Load balancing tools are used to direct incoming transactions to the server in the server pool in such a way that the traffic is balanced across all the servers in the pool. As such, the transactions can be processed faster and more efficiently.

[0004] One approach to load balancing simply involves routing each new transaction to a next server in the server pool (i.e., the "round-robin" approach). However, this approach does not distinguish between available servers and those which are down or otherwise unavailable. Therefore, transactions directed to unavailable servers are not processed in a timely manner, if at all. Other approaches to load balancing involve routing transactions to the next available server. That is, an agent monitors a pool of servers for failure and tags servers that are unavailable so that the load balancer does not route transactions to an unavailable server. However, this approach is also inefficient, still not necessarily routing transactions to the server that is best able to process the transaction. For example, a large transaction (e.g., a video clip) may be directed to a slow server even though there is a faster server available, because the slow server is identified as being the "next available" server when the transaction arrives at the load balancer. Likewise, a low priority transaction (e.g., an email) may be directed to the fast server simply based on the order that the servers become or are considered available.

[0005] A more current approach uses a combination of system-level metrics to route transactions and thus more efficiently balance the incoming load. The most common metrics are based on network proximity. For example, the 3/DNS load balancing product (available from F5 Networks, Inc., Seattle, Wash.) probes the servers and measures the packet rate, Web-request completion rate, round-trip time and network topology information. Also for example, the Resonate Global Dispatch load balancing product (available from Resonate, Inc., Sunnyvale, California) uses latency measurements for load balancing decisions.

[0006] However, while system metric approaches measure server characteristics, the transaction is not routed based on

service levels required by or otherwise specific to the transaction. That is, the transaction is not routed based on the transaction size, the originating application, the priority of the transaction, the identification of the user generating the transaction, etc. Instead, the transaction is routed to the fastest available server when the transaction arrives at the load balancer. As such, the video clip and the low priority email, in the example given above, still may not be efficiently routed to the servers for processing. For example, if the low priority email arrives at the load balancer when the fastest server is available, the email will be routed to the fastest server, thus leaving only slower servers available when the high priority video clip later arrives at the load balancer.

[0007] Likewise, transactions are often directed to other network devices (e.g., routers, servers, storage devices, etc.) based merely on an IP address. Again, such an approach may not be the most efficient when routing transactions from different applications, users, at different times, etc.

SUMMARY OF THE INVENTION

[0008] The inventors have devised a method and apparatus for identifying a requested level of service for a transaction wherein the transaction may be processed in accordance with the requested level of service.

[0009] The invention is preferably embodied in computer readable program code stored in suitable storage media, and comprises program code for selecting the requested level of service for the transaction, and program code for assigning the requested level of service to the transaction. Preferably, the transaction is a packetized signal having at least a data packet (e.g., the data to be processed), and a service tag including the requested level of service.

[0010] ~~The requested level of service can be based on any suitable factors, and can be assigned at any point on the network. For example, the service tag can indicate the requested level of service as a predefined service category (e.g., premium, standard, low), a user identification (e.g., user1, user2, administrator), a transaction type (e.g., email, video), etc. Also for example, the service tag can be user-defined, set by the application submitting the transaction, set by an administrator, based on the time (e.g., weekday or weekend), based on the type of transaction, etc. Or for example, the requested level of service may be user-defined (e.g., using commands or strings of commands) via a suitable user interface. Or the requested level of service may be based on a combination of factors. For example, a transaction may include a requested level of service and a backup level of service, wherein the transaction is directed to the network device based on the backup level of service when the requested level of service is unavailable.~~

[0011] The service tag is read from the transaction using suitable program code (e.g., at a load balancer). Based on the requested level of service, the transaction is directed to a network device (e.g., a server) that is best able to provide the requested level of service for processing the transaction. For example, where the requested level of service associated with the transaction is a scale value of "50", the load balancer selects the server providing a corresponding service level nearest the requested level of service, such as a scale value of "48". Alternatively, the transaction can be processed by a server within a group of servers wherein each

server best provides the requested level of service. For example, a category of service can be requested, such as "premium", and the load balancer thus selects any server from the group of servers providing a corresponding service level of "premium".

[0012] As such, the transaction is efficiently routed to a server based on service level information specific to the transaction. Thus for example, a low priority transaction (e.g., an email) may arrive at the load balancer before a high priority transaction (e.g., a video clip) when the fastest server is available. However, the low priority transaction is identified as such and routed to a slower server. Thus, the fastest server is available when the high priority transaction arrives at the load balancer, even so it arrives later than the low priority transaction.

[0013] These and other important advantages and objectives of the present invention will be further explained in, or will become apparent from, the accompanying description, drawings and claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014] Illustrative and presently preferred embodiments of the invention are illustrated in the drawings in which:

[0015] FIG. 1 shows a first embodiment of a load balancer for routing a transaction to a server;

[0016] FIG. 2 shows a packetized transaction having a service tag associated therewith for requesting a level of service for the transaction;

[0017] FIG. 3 shows a second embodiment of a load balancer for routing the transaction of FIG. 2 to a server based on the requested level of service indicated by the service tag;

[0018] FIG. 4 illustrates a server index identifying servers and the corresponding service level of each server that can be used by the load balancer in FIG. 3;

[0019] FIG. 5 shows a load balancer routing the transaction of FIG. 2 to a server within a group of servers each best able to provide the requested level of service indicated by the service tag;

[0020] FIG. 6 illustrates a server index identifying groups of servers and the corresponding service level of each group that can be used by the load balancer in FIG. 5;

[0021] FIG. 7 is a flow chart showing a method for routing the transaction of FIG. 2 to a server, as in FIG. 3 and FIG. 5;

[0022] FIG. 8 is a flow chart showing a method for identifying a requested level of service for a transaction for processing in accordance with the requested level of service; and

[0023] FIG. 9 illustrates various points of a system where the service tag may be assigned to the transaction.

DESCRIPTION OF THE PREFERRED EMBODIMENT

[0024] FIG. 1 shows a load balancer 100 for routing a transaction 110 to a number of (i.e., one or more) servers 121, 122, 123 in a server pool 120. For purposes of illustration, Server A is unavailable as indicated by the "X" in

FIG. 1. Using a simple "round-robin" approach, the load balancer 100 receives a next transaction 110 and directs the transaction 110 to the next server in the server pool 120 (i.e., the last server to have received a transaction). For example, where the previous transaction is directed to server 123 (Server C), the next server is server 121 (Server A) even where the server 121 (Server A) is unavailable as shown in FIG. 1, and so forth. Alternatively, the load balancer 100 directs the transaction 110 to the next available server in the server pool 120. That is, an agent (e.g., suitable program code) monitors each of the servers 121, 122, 123 in the server pool 120 and labels a server that has failed, shut down, or is otherwise unavailable, as "unavailable" (e.g., using a suitable computer readable tag). Thus, the load balancer 100 recognizes a server that has been labeled "unavailable" and does not route transactions to the unavailable server. For example, where the previous transaction was directed to server 123 (Server C) and server 121 (Server A) is indicated as being "unavailable", the next server is server 121 (Server A). However, the next available server is server 122 (Server B). Therefore, in this example the transaction 110 is directed to server 122 (Server B). Alternatively, the load balancer 100 can direct the transaction 110 to the "fastest" available server in the server pool 120. For example, where server 121 (Server A) generally provides a fast turn-around but is labeled "unavailable", server 122 (Server B) provides a medium turn-around, and server 123 (Server C) provides a slow turn-around, the transaction 110 is routed to server 122 (Server B). That is, although server 121 (Server A) is generally the fastest server in the server pool 120, server 121 (Server A) is unavailable, therefore leaving server 122 (Server B) as the fastest available server. However, none of these approaches direct the transaction 110 to a server 121, 122, 123 based on parameters specific to the transaction 110.

[0025] FIG. 2 shows a packetized transaction 200. The packetized transaction 200 includes a data packet 210 (i.e., the data to be processed) and a service tag 220. The data packet 210 can include any data that is to be processed in any number of ways. For example, the data packet 210 can include an email message to be delivered to a recipient, a uniform resource locator (URL) requesting a hypertext markup language (HTML) page from the corresponding Internet site, data to be stored in a network area storage (NAS) device, spreadsheet data for tabulation, a portion thereof to be reassembled upon reaching the network device, etc. Optionally, the transaction 200 can also include additional fields. For example, the transaction 200 can include a destination field 230, such as a device on the network indicated or identified by an IP address.

[0026] The service tag 220 is preferably a single or multi-bit packet associated with the data packet 210, the value of which identifies a requested level of service for the transaction 200. However, the service tag 220 and the requested service level identified therein can take any suitable form. The service tag 220 can be a numeric value. For example, the service tag 220 may be a single bit such as a "one", indicating high priority, or a "zero", indicating low priority. Or for example, the service tag can be a number on a scale of one to ten, with each number indicating a level of service. Alternatively, the service tag 220 can indicate the requested level of service as a predefined service category. For example, the requested level of service can be indicated as premium, standard, or low. Likewise, the service tag 220

can indicate the requested level of service as a specific parameter. For example, the service tag 220 can indicate the processing speed or capacity required or desired for the transaction 200. Or for example, the service tag 220 can indicate a turn-around time for the transaction.

[0027] It is also understood that the service tag 220 can include multiple packets. Similarly, an individual service tag 220 may comprise more than one indicator. These multiple packets, or indicators within a packet, may be combined to indicate the requested level of service. Separate packets (or indicators) may be included, such as, a time-stamp, an origination ID, an application ID, a user ID, a project ID, etc. In such an embodiment, the requested service level can be a combination of some or all of the packets included therein. For example, a transaction 200 can be handled with high priority where the user ID is "administrator" and the origination ID is "network administration tools". Or for example, a transaction 200 may be handled with high priority where the user ID is "user 1" and the time-stamp indicates the transaction must be completed during business hours. Likewise, a transaction 200 also having a user ID of "user 1" may be handled with lower priority where the time-stamp indicates that the transaction can be completed during off-peak hours.

[0028] Alternatively, these multiple packets can individually indicate or identify the requested level of service. For example, the service tag 220 may indicate a preferred level of service and a backup level of service. In such an embodiment, where the preferred level of service is unavailable, the transaction is preferably handled or processed according to the backup level of service. In addition, in this embodiment the transaction 200 may be "bounced" (i.e., returned to the originating computer) where neither the preferred level of service nor the backup level of service can be provided. Or a warning or a message may be returned to the originator where either the backup level of service is provided or where neither is provided.

[0029] In yet another alternative, the transaction 200 can be prioritized based on a first packet (or indicator) or set of packets (or indicators), and the remaining packets (or indicators) are read or factored in only where there is a conflict between service levels requested by more than one transaction 200. For example, where the user ID packet is "administrator" for more than one transaction 200, the competing transactions 200 can be hierarchically prioritized based on the application ID packet, and so forth.

[0030] It is also understood that the requested level of service need not be a separate packet (i.e., service tag 220) associated with data packet 200. For example, the requested level of service can be included as part of the destination packet 230. Or for example, the requested level of service may be included with the data packet 210.

[0031] Preferably, the transaction 200 is assigned a service tag 220 at its source (i.e., where the transaction 200 originates). However, it is understood that the service tag 220 may be assigned and/or changed at any suitable device along the transaction path, as explained below with respect to FIG. 9.

[0032] It is understood that the requested level of service indicated by the service tag 220 may be assigned to the transaction 200 based on any number of factors. The

requested level of service may be based on the time sensitivity of the transaction 200. For example, data that is time sensitive may be assigned a higher priority than data that is not time sensitive. Or for example, a transaction 200 that would normally be assigned to a slow server during business hours can be assigned to a faster server during evening hours and on weekends. The requested level of service may also be based on characteristics or parameters of the transaction 200 itself. For example, large processing requests can be assigned to faster servers. The requested level of service may also be based on the user identification. For example, users that generally require faster processing speeds (the CAD department or an administrator) may be assigned faster servers than those who require the servers only to back up data. Likewise, users (e.g., an administrator) can be designated as having the highest priority, overriding competing transactions.

[0033] In addition, the requested level of service may be manually assigned to a transaction 200. That is, the requested level of service may be user-defined. Suitable program code (e.g., a user interface) can be included at a network device (e.g., a workstation, router, etc.) or as part of an application (e.g., an operating system, the originating application, an applet, etc.). For example, the user may be queried before starting an application as to the requested level of service. The user might have predefined options (e.g., a menu, list, etc.). Or the user may specify a requested level of service (e.g., using predefined commands or command strings). Alternatively, the requested level of service may be set by an administrator. For example, the administrator may assign a priority to a particular workgroup, user, project, application, etc.

[0034] In another embodiment, the requested level of service may be automatically assigned to a transaction 200. That is, the requested level of service may be assigned by a network device (e.g., a workstation, router, etc.) or an application (e.g., an operating system, the originating application, an applet, etc.). Suitable program code can be included at the network device and/or the application to assign the requested level of service based on a fixed parameter (e.g., a user ID, application ID, passcode, etc.), a dynamic parameter (e.g., the time of day, current load, etc.) or a combination thereof.

[0035] In yet another embodiment, the requested level of service may be selected through a combination of manual and automatic identification. For example, the user may define a requested level of service, and the transaction may be automatically marked with an application ID, a time-stamp, etc. In such an embodiment, the transaction 200 may be handled based on the user-defined level of service, the automated markings, or a combination thereof.

[0036] It is understood that the above examples are merely illustrative of the requested level of service and associating it with the data packet 210 (e.g., assigned to the transaction 200). Other examples are also contemplated as within the scope of the present invention.

[0037] FIG. 3 shows the transaction 200 received at a load balancer 300 and directed to a server 311, 312, 313 in a server pool 310 that is best able to process the transaction 200 based on the requested level of service indicated by the service tag 220. In FIG. 3, the load balancer 300 selected

server 312 (Server B) as the server that is best able to process the transaction 200, using the service tag 220 and the server index 400 (FIG. 4).

[0038] The server index 400 (FIG. 4) is preferably a multi-dimensional array (e.g., a database or "lookup table") stored in a memory accessible by the load balancer 300. The server index 400 includes at least a server identification (ID) 410 and a corresponding service level 420 for each server 311, 312, 313 in the server pool 320 that is managed by the load balancer 300. The server ID 410 can be the server IP address, a path, or any other suitable means that the load balancer 300 can use to identify a server 311, 312, 313 and direct a transaction 200 thereto. Other data related to the various servers can also be included in the server index, such as that status of a particular server (e.g., availability, current load), alternative or backup servers or server pools, etc.

[0039] The service level 420 can be any suitable indicator, such as but not limited to a number on a scale of one to ten, a category of service, the time (e.g., weekday or weekend), a user identification (e.g., user1, user2, administrator), a transaction type (e.g., email, video), a combination thereof, etc. Furthermore, the service level can be based on information about the monitored servers obtained by polling the servers, predefined service specifications, etc. Likewise, the servers can be ranked relative to one another, relative to the types of transactions processed, etc.

[0040] When the transaction 200 is received by the load balancer 300, the service tag 220 is read using suitable program code. The load balancer 300 then accesses the server index 400 to determine (e.g., using suitable program code) the server in the server pool 310 that can best provide the requested level of service associated with the transaction 200 (i.e., as indicated by the service tag 220). For example, where the service tag 220 indicates a requested level of service having a scale value of "50", the server index 400 indicates that server 312 (Server B) is providing a corresponding service level 420 having a scaled value of "51", while the other servers 311 and 313 are providing lower levels of service. Hence, the load balancer 300 directs the transaction to server 311 (Server B), as shown in FIG. 3. As another example, where the service tag 220 indicates the requested level of service is a scaled value of "25", the load balancer 300 directs the transaction 200 to server 313 (Server C), which is providing a corresponding service level 420 having a scaled value of "27", as indicated by the server index 400.

[0041] It is to be understood that the term "best", as that term is used herein with respect to the server best able to provide the requested level of service, is defined to mean "best as determined by the program code of the load balancer", and may be interpreted by a load balancer as, for example, "nearest" or "meeting" the requested level of service. Thus, even where the requested level of service and the service level actually being provided are at opposite ends of a spectrum (e.g., the requested level of service is a scaled value of "50" but the service levels being provided by the servers range from scaled values of "5" to "10"), the server providing the service level nearest to that requested (e.g., a service level having a scaled value of "10") is considered to be "best" able to provide the requested level of service. However, it is also to be understood that where the disparity between the requested level of service and the service level

being provided is unacceptable (i.e., based on a predetermined level of acceptability, such as more than "10" scale values difference), the load balancer 300 can direct the transaction to the server best able to provide the requested service level, but also return a warning signal (e.g., an email, an error message, etc.) to the requester (e.g., an administrator, the user, the originating application, etc.) notifying the requestor of the disparity. Alternatively, the load balancer 300 can redirect the transaction 200 to another load balancer that is monitoring another pool of servers, the load balancer 300 can "bounce" the transaction 200 altogether, etc.

[0042] It is also to be understood that the term "server" as used herein can be any computer or device that manages resources, such as a file server, a printer server, a network server, a database server, etc. In addition, the servers can be dedicated or the servers can be partitioned (i.e., have multiprocessing capability), in which case the term "server" may instead refer to software that is managing resources rather than to an entire computer or other hardware device.

[0043] In FIG. 5, the server pool 500 includes a premium group 510, a standard group 520, and a low priority group 530. The servers 511, 512, and 513 (A, B, and C, respectively) are part of the "premium" group 510. For example, the premium group 510 can include high-speed, high-capacity servers. In addition, the premium group 510 can include additional servers and backup servers so that there is always an available server in this group. Access to these servers can be reserved for a department with high demand requirements (e.g., the CAD department), for high priority transactions, for customers paying a fee to access these servers, etc. The standard group 520 can include average-speed, average capacity servers. Access to these servers 521, 522 (D and E) can be designated for a sales/marketing department that requires only average processing capacity, or can also be available on a fee-basis. The "low priority" group 530 can include older and/or less expensive servers 531 that do not perform at the predetermined standards of the standard group 520 or the premium group 510. These servers 531 can be used for low-priority email, backup jobs, transactions requested during off-peak hours when timeliness is not as important, etc. These servers can be designated as a group 530, or simply be unclassified servers in the server pool 500.

[0044] It is to be understood that any number of groups can be designated. The manner in which groups are designated can include static parameters such as processing speed, capacity, server proximity, etc. However, preferably the groups 510, 520, 530 are dynamically designated based on monitored performance of the individual servers. For example, where a "premium" server (e.g., 511) is not performing to a predetermined standard, it can be reclassified as a standard or low priority server (i.e., in group 530), whereas a standard server (e.g., 521) that has recently been upgraded can be reclassified as a premium server (i.e., in group 510). Likewise, the invention disclosed herein is not to be limited by the groups 510, 520, 530 shown in FIG. 5. For example, more or fewer groups can be used, servers can be further subdivided within the groups, the groups can be identified by means other than the labels "premium", "standard", and "low", etc.

[0045] The service level being provided by each server can be based on, as illustrative but not limited to, the server meeting the service level objectives of a single user, a user

group (e.g., the accounting department), or a transaction type (e.g., email). That is, preferably the load balancer 300 (or suitable software/hardware agent) monitors the service level provided by each server in the server pool to generate the server index. For example, the load balancer 300 can measure or track processing parameters of a server (e.g., total processing time, processor speed for various transactions, etc.) with respect to a single user, a user group, a transaction type, etc. Alternatively, the server index can be based on known capabilities (e.g., processor speed, memory capacity, etc.) and/or predicted service levels of the servers in the server pool (e.g., based on past performance, server specifications, etc.). Or for example, the load balancer 300 can access multiple server indexes, wherein each index is based on a different set of monitored server parameters. A group ID or the like associated with a transaction can then be used as the basis for the load balancer 300 accessing a particular server index.

[0046] In any event, it is understood that the service level provided by each server in the server pool can be formatted similar to the requested level of service. Alternatively, program code for translation can be implemented (e.g., at the load balancer 300) to convert between formats. For example, a category of service level, such as "premium", associated with the transaction 200 can be converted to a scale value, such as "50", associated with a server or group of servers in the server pool.

[0047] When the transaction 200 is received at the load balancer 300, the load balancer 300 reads the requested level of service from the service tag 220. Based on the server index 600 (FIG. 6), the load balancer 300 selects the server (e.g., 512) from the server group (e.g., 510) that is best providing the requested level of service (e.g., "premium"). That is, the server index 600 contains the server ID 610 and a corresponding level of service 620, similar to the server index 400 in FIG. 4. However, in server index 600, the server ID 610 is indicated as a group of servers. That is, Servers A, B, and C, are providing a "premium" level of service, Servers D and E are providing a "standard" level of service, and Server F is providing a low-priority level of service. Thus for example, where the service tag 220 indicates that the requested level of service is "premium", the load balancer 300 directs the transaction 200 to any one of the servers 511, 512, 513 in the premium group 510. The load balancer can use conventional load balancing algorithms (e.g., next available, fastest available, or any other suitable algorithm) to select a specific server 511, 512, 513 within the premium group 510.

[0048] It is understood that the load balancing schemes shown in FIG. 3 and FIG. 5 are illustrative of the apparatus and method of the present invention and are not intended to limit the scope of the invention. Other configurations are also contemplated as being within the scope of the invention. For example, multiple load balancers can be networked to administer a single server pool or multiple server pools. Such a configuration allows a load balancer experiencing heavy use to transfer some or all of the transactions in bulk to another load balancer experiencing a lighter load. Or for example, a hierarchy of load balancers might administer the server pool. A possible hierarchical configuration could comprise a gatekeeping load balancer that directs transactions either to a load balancer monitoring a premium server pool or to a load balancer monitoring a standard server pool, and the individual load balancers can then select a server from within the respective server pool.

[0049] FIG. 7 shows a method for routing the transaction 200 to a server based on a requested level of service associated with the transaction 200 generated in step 710, using suitable program code and stored on a number of (i.e., one or more) suitable computer readable storage media. In step 700, the load balancer 300 (or a suitable software/hardware agent) monitors the server pool 320, 500 to determine the service level of each server in the server pool. In step 710, the load balancer 300 (or a suitable software agent) uses the monitored data to generate a server index (e.g., 400, 600) having at least the server ID (e.g., 410, 610) and the corresponding service level (e.g., 420, 620), including groups of servers where desired. In step 720, when a transaction 200 is received at the load balancer 300, the load balancer 300 (or suitable program code associated therewith) reads the requested level of service indicated by the service tag 220 associated with the transaction 200. In step 730, the load balancer 300 accesses the server index to select a server from the server pool that is best able to provide the requested level of service. Once a server has been selected, the load balancer 300 directs the transaction 200 to the selected server in the server pool in step 740.

[0050] It is understood that the method shown and described with respect to FIG. 7 is merely illustrative of a preferred embodiment. However, each step need not be performed under the teachings of the present invention. Step 710 can be modified or eliminated, as an example, where a server index is provided with a predetermined server ID and the corresponding service level is packaged with the load balancer 300. Likewise, the steps need not be performed in the order shown in FIG. 7. For example, the transaction 200 can be received and the service tag 220 read by the load balancer (as in step 720), followed by the load balancer 300 monitoring the server pool for a server providing the requested level of service (as in step 700). In such an example, it is also understood that a server index need not be generated at all (as in step 710) and that the load balancer can select a server dynamically (i.e., based on current server performance).

[0051] FIG. 8 shows a method for identifying a requested level of service for a transaction, wherein the transaction may be processed in accordance with the requested level of service. In step 800, the requested level of service is selected for the transaction 200. The requested level of service can be selected by the original application, an administrator, etc., as explained above. Likewise, the requested level of service can be based on any suitable factors and assigned by any suitable device on the network. In step 810, the requested level of service is assigned to the transaction 200 (e.g., as a service tag 220 associated with the data packet 210). In step 820, the service tag 220 is read (e.g., by a load balancer 300 in FIG. 3 and FIG. 5) using suitable program code. In step 830, the transaction 200 is directed to and processed by a network device, such as a server in server pool 310, 500 (FIG. 3 and FIG. 5, respectively) based on the service tag 220, as discussed above.

[0052] FIG. 9 illustrates various points of a system where the service tag 220 may be assigned to the transaction 200. For example, the service tag 220 may be assigned to the transaction 200 at the originating application 900 via a graphical user interface (GUI) 905 or via other suitable program code 910 integrated as part of the application or interacting therewith. Or for example, the service tag 220 may be assigned at a workstation 920 via the operating system (OS) 925 or other suitable program code executed on the workstation 920. In another example, the service tag 220

may be assigned on the local area network 930 by a network component 935 (e.g., a server, another workstation, a router, hub, etc.), and/or on the wide area network 940, also by a network component 945.

[0053] It is understood that the service tag 220 may be assigned to the transaction 200 at any number of points, and FIG. 9 is merely intended to be an illustration thereof. Other examples include assigning the service tag 220 to the transaction 200 by an intermediary computer, a gateway, a load balancer, etc. Another example may include dynamically assigning the service tag 220. That is, the service tag 220 may be assigned to the transaction 200 at the originating application 910, then appended at the operating system 925, and further changed at a router 935 and/or 945 on the network 930 and/or 940 before the transaction 200 reaches the destination. More specifically, the user may request a level of service of "high priority" for a transaction 200 via GUI 905. The operating system 925 may subsequently append a backup level of service of "medium priority" to the transaction 200. A router 935 and/or 945 that receives the transaction 200 while handling a heavy load from a high priority user (e.g., an administrator), may then change the requested level of service to "best available". As such, the service tag 220 need not be statically assigned.

[0054] While illustrative and presently preferred embodiments of the invention have been described in detail herein, it is to be understood that the inventive concepts may be otherwise variously embodied and employed, and that the appended claims are intended to be construed to include such variations, except as limited by the prior art.

What is claimed is:

1. An apparatus for identifying a requested level of service for a transaction, comprising:

computer readable storage media; and

computer readable program code stored in said storage media, comprising:

- a) program code for selecting said requested level of service for said transaction; and
- b) program code for assigning said requested level of service to said transaction.

2. An apparatus, as in claim 1, wherein said transaction is a packetized signal comprising at least a data packet, and wherein a service tag is associated with said data packet by said program code for assigning said requested level of service, said service tag including said requested level of service.

3. An apparatus, as in claim 1, further comprising:

- a) program code for selecting a backup level of service; and
- b) program code for assigning said backup level of service to said transaction.

4. An apparatus, as in claim 1, wherein said requested level of service is a predefined service category.

5. An apparatus, as in claim 1, wherein said requested level of service is based on a user identification.

6. An apparatus, as in claim 1, wherein said requested level of service is based on a transaction type.

7. An apparatus, as in claim 1, further comprising a user interface for selecting said requested level of service.

8. An apparatus, as in claim 1, wherein said requested level of service includes a plurality of parameters.

9. A method for requesting a level of service for a transaction on a network, comprising:

selecting said requested level of service for said transaction; and

assigning said requested level of service to said transaction.

10. A method, as in claim 9, wherein selecting said requested level of service comprises receiving a user-defined level of service.

11. A method, as in claim 9, wherein selecting said requested level of service comprises assessing a number of characteristics of said transaction.

12. A method, as in claim 9, wherein a network device best provides said requested level of service.

13. A method, as in claim 9, wherein said requested level of service is automatically assigned to said transaction.

14. An apparatus for routing a transaction over a network based on a requested level of service associated with said transaction, comprising:

a number of computer readable storage media; and

computer readable program code stored in said number of storage media, comprising:

- a) program code for selecting said requested level of service for said transaction;
- b) program code for assigning a service tag to said transaction, said service tag including said requested level of service;
- c) program code for reading said requested level of service from said service tag; and
- d) program code for directing said transaction over said network based on said requested level of service read from said service tag.

15. An apparatus, as in claim 14, wherein said transaction is directed over said network to a device best providing said requested level of service.

16. An apparatus, as in claim 14, wherein said service tag is assigned by program code at more than one point on said network.

17. An apparatus, as in claim 14, wherein said service tag is read by program code at more than one point on said network.

18. An apparatus, as in claim 14, further comprising program code for changing said requested level of service included on said service tag.

19. An apparatus for requesting a level of service for a transaction on a network, comprising:

means for selecting said requested level of service; and

means for assigning said requested level of service to said transaction, wherein said transaction is directed to a network device based on said requested level of service.

20. An apparatus, as in claim 19, further comprising means for reading said requested level of service assigned to said transaction.

* * * * *



US006052715A

United States Patent [19]**Fukui et al.**[11] **Patent Number:** **6,052,715**[45] **Date of Patent:** **Apr. 18, 2000**[54] **INTERACTIVE COMMUNICATION SYSTEM
FOR DOWNLOADING LARGE AMOUNT
DATA**[75] **Inventors:** Takayoshi Fukui, Hamura; Masayasu
Miyake, Fuchu, both of Japan[73] **Assignee:** Casio Computer Co., Ltd., Tokyo,
Japan[21] **Appl. No.:** 08/855,455[22] **Filed:** May 13, 1997[30] **Foreign Application Priority Data**

May 16, 1996 [JP] Japan 8-122057

[51] **Int. Cl.⁷** H04H 1/00; H04N 7/173[52] **U.S. Cl.** 709/217; 348/12[58] **Field of Search** 348/6, 7, 10, 12,
348/13; 455/3.1, 3.2, 4.1, 4.2, 6.1, 6.2,
5.1; 709/217-219[56] **References Cited****U.S. PATENT DOCUMENTS**

5,177,604	1/1993	Martinez .	
5,321,514	6/1994	Martinez .	
5,347,304	9/1994	Moura et al.	348/12
5,761,602	6/1998	Wagner et al.	455/3.1

Primary Examiner—Nathan Flynn
Attorney, Agent, or Firm—Frishauf, Holtz, Goodman,
Langer & Chick, P.C.[57] **ABSTRACT**

An information terminal transmits small amount data such as a command to a Web server on the internet to an asymmetric router via an interactive radio network. The asymmetric router transmits the command to the Web server via the internet. The asymmetric router transmits large amount data (for example, HTML formatted browser data) from the Web server to a data superimposing broadcast apparatus. The data superimposing broadcast apparatus determines a scheduled time of broadcasting according to a data amount. The scheduled time is transmitted to the information terminal via the asymmetric router and the interactive radio network. When the scheduled time comes, the large amount data is superimposed on an empty portion of a television signal, and is broadcasted. The information terminal carries out a predetermined processing operation whereby a broadcast receiving circuit is turned on, and the television signal on which the large amount data has been superimposed is received. The received data then can be displayed in a display device and the like. Therefore, even in a portable information terminal, it is possible to provide an interactive communication system which can transmit large amount data at a high speed.

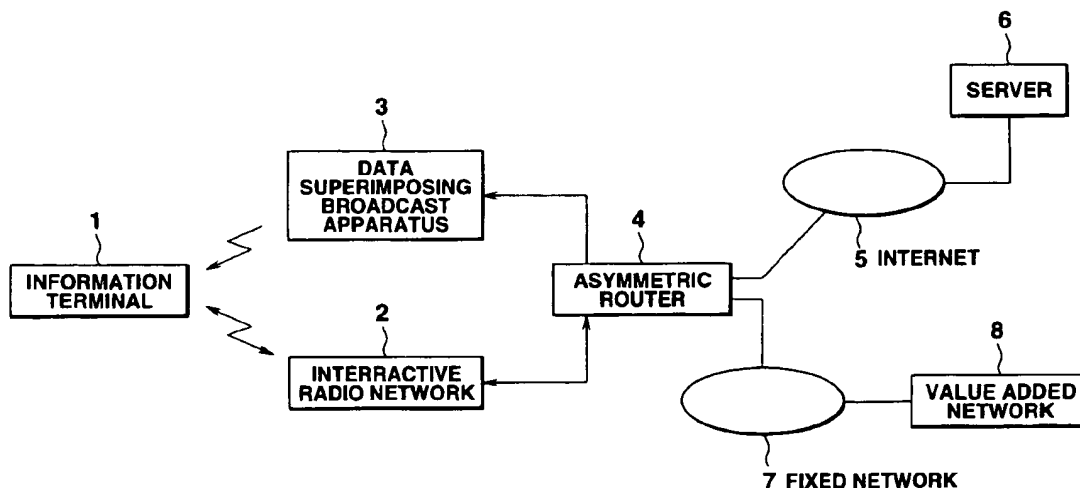
13 Claims, 14 Drawing Sheets

FIG. 1

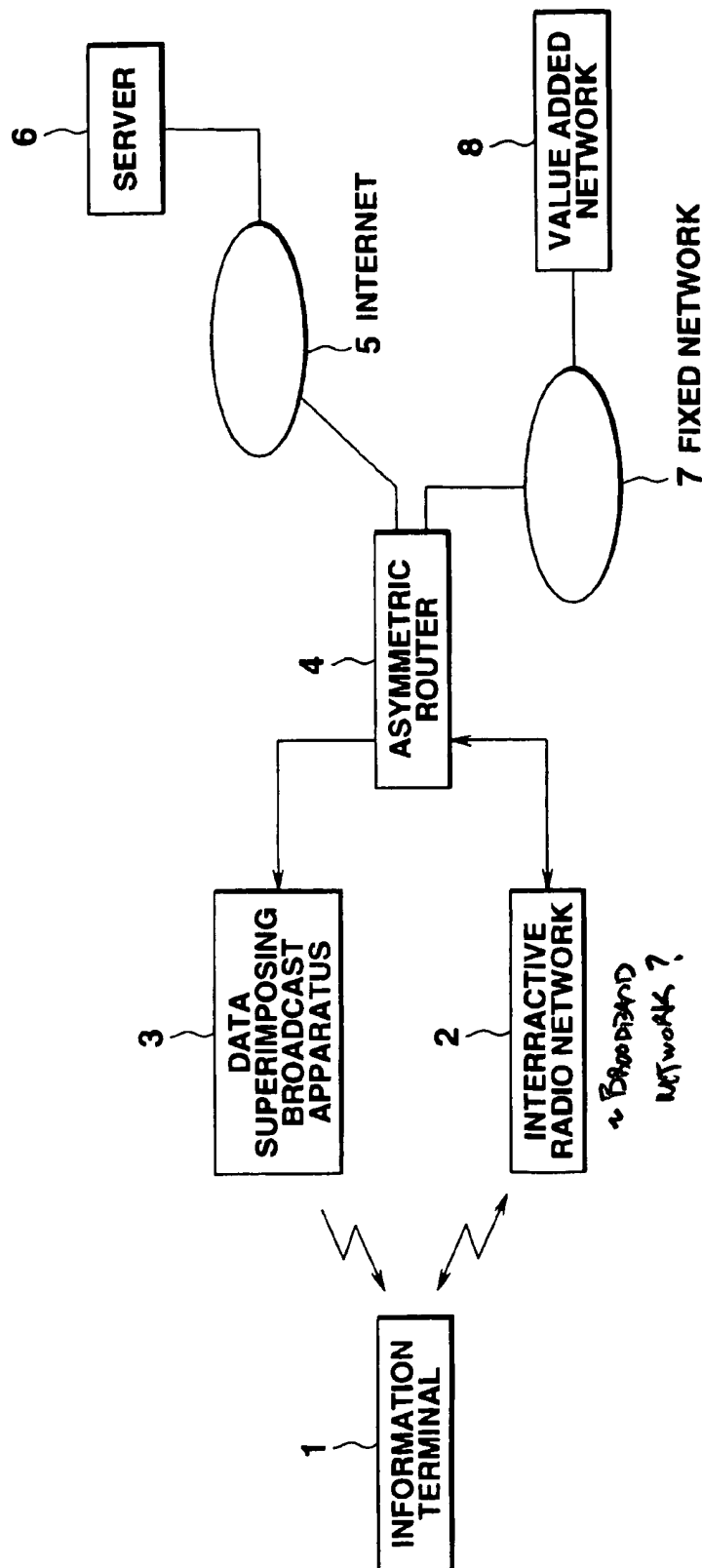


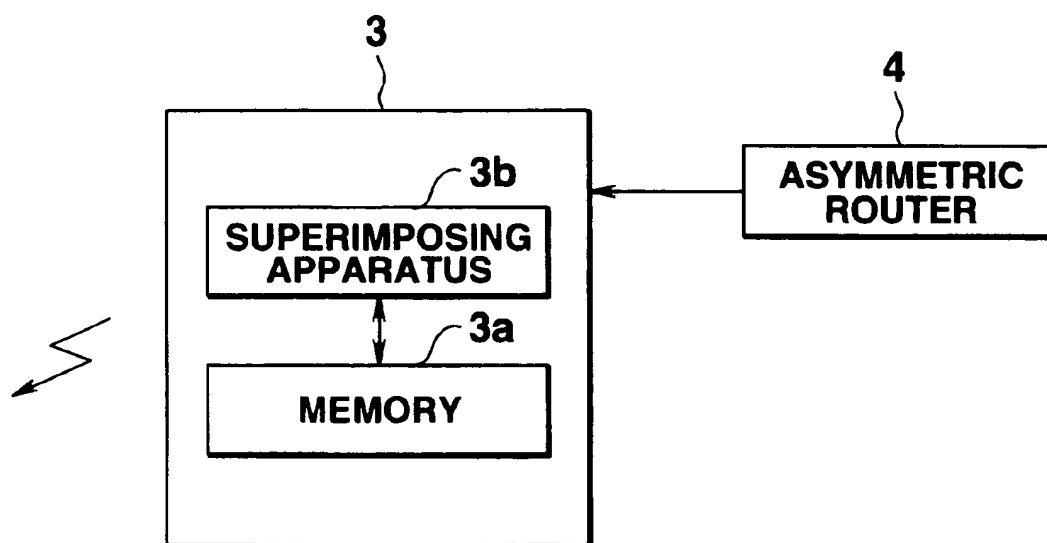
FIG.2

FIG.3

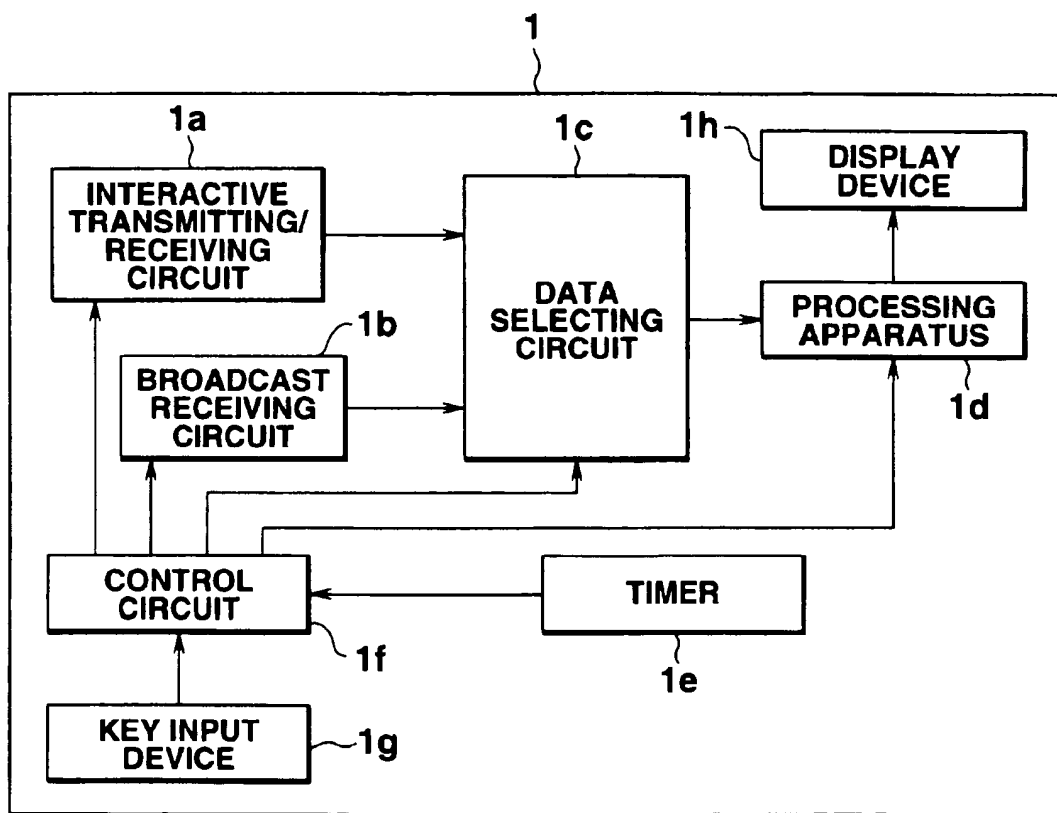


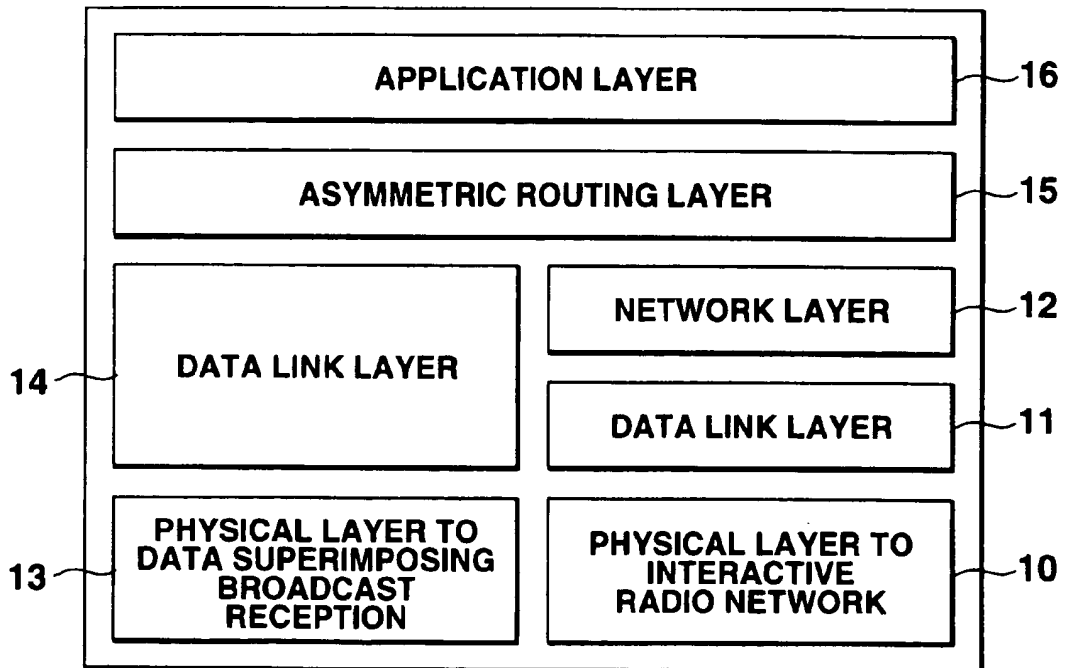
FIG.4

FIG. 5

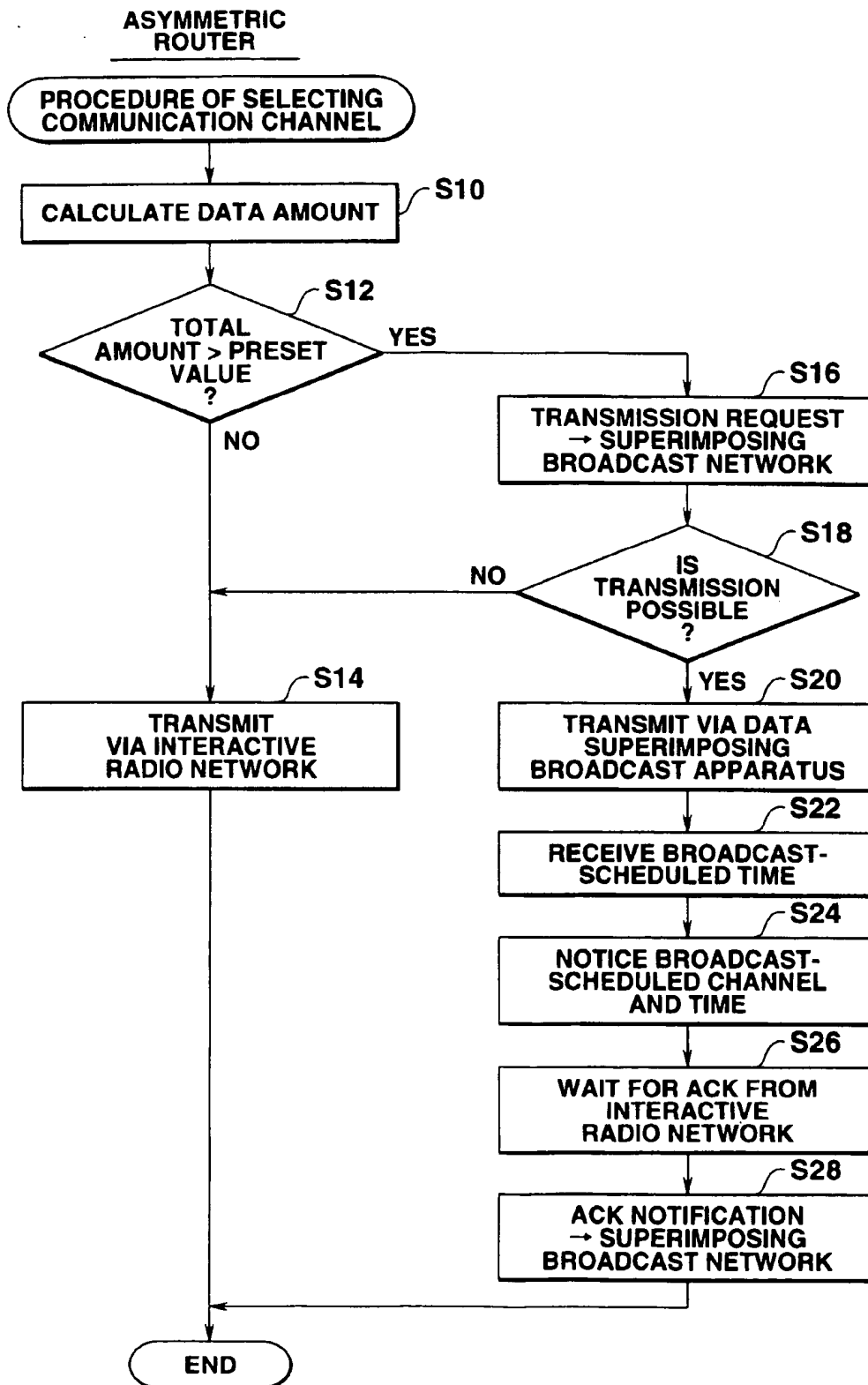


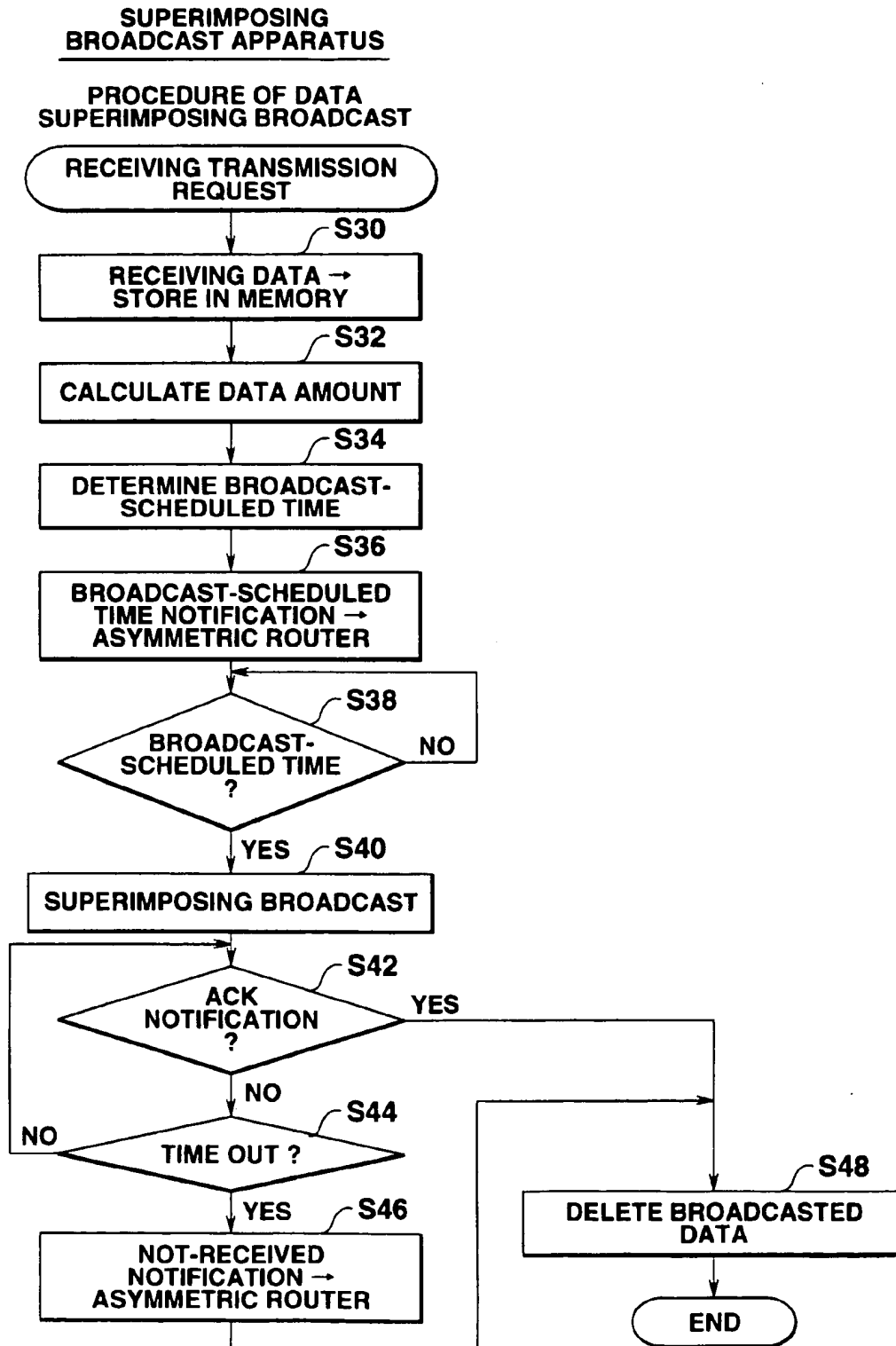
FIG.6

FIG. 7

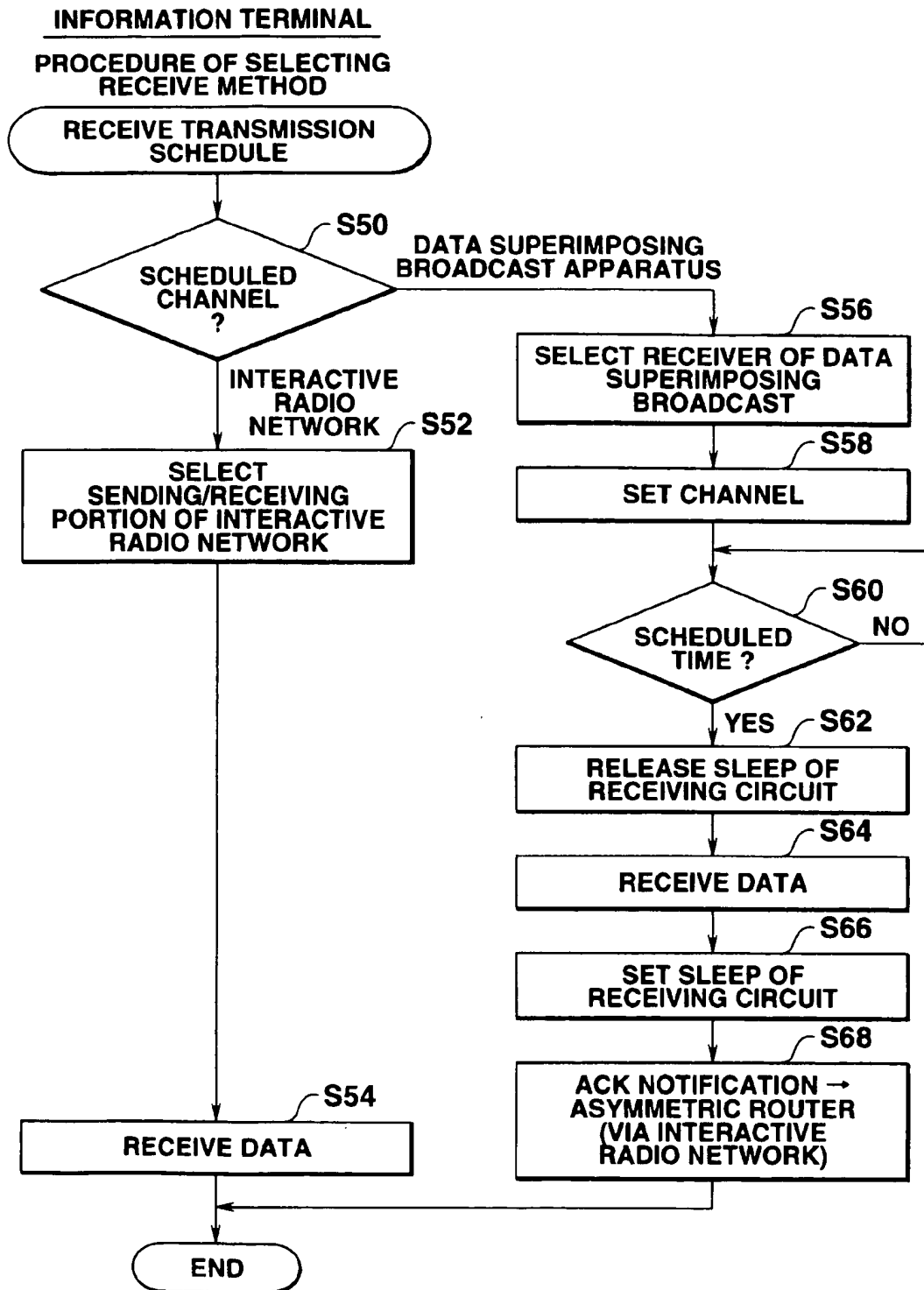


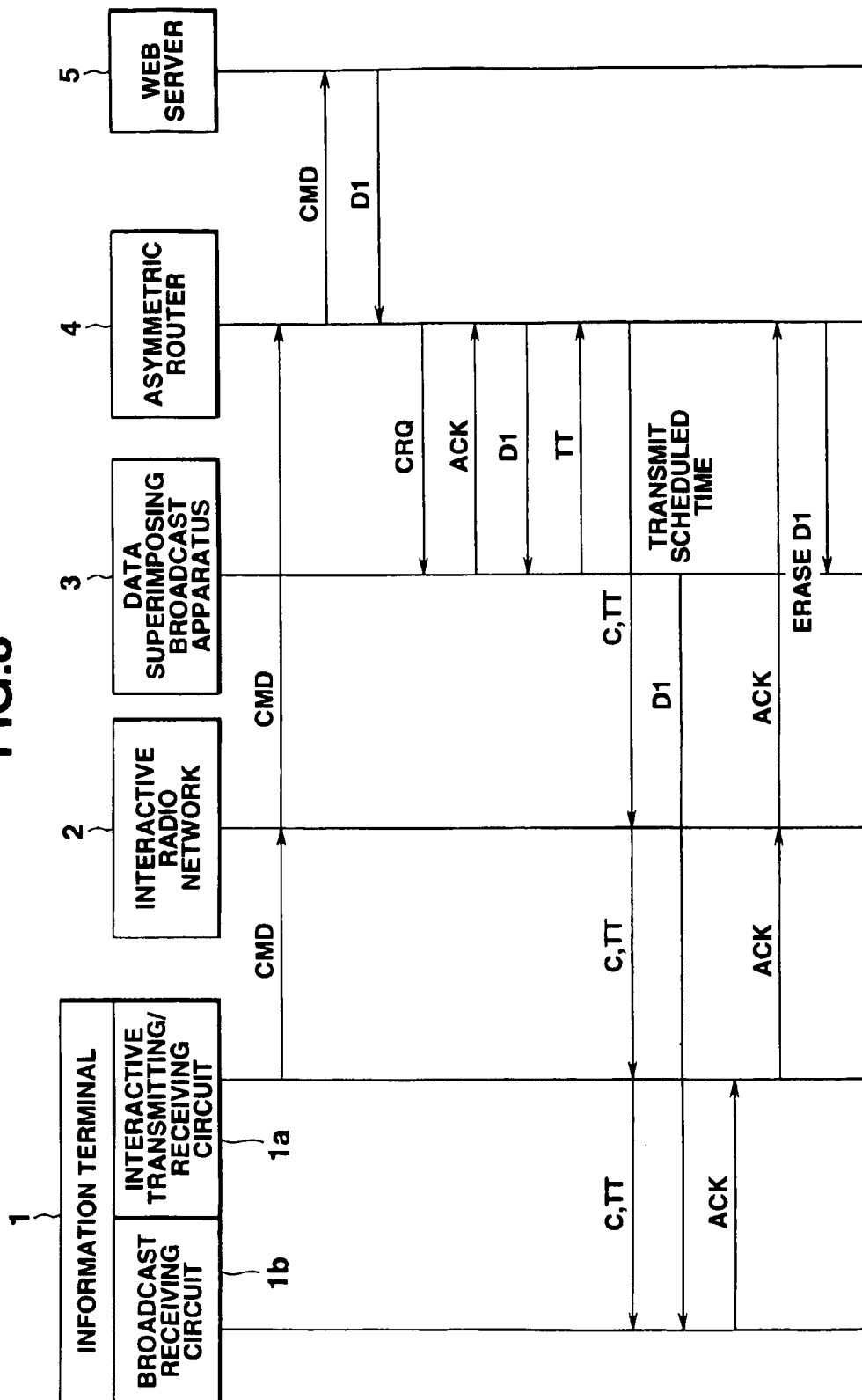
FIG.8

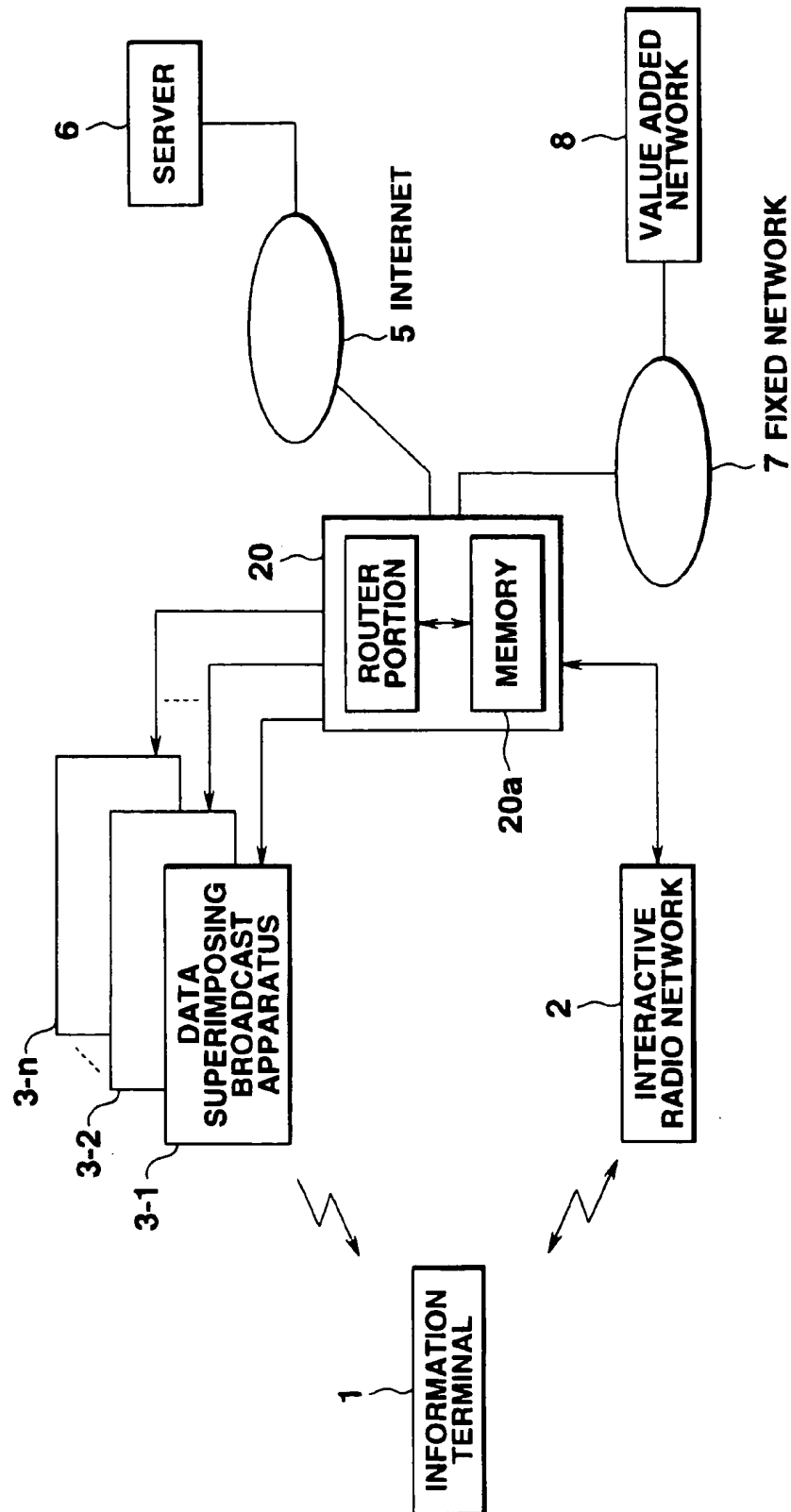
FIG. 9

FIG. 10

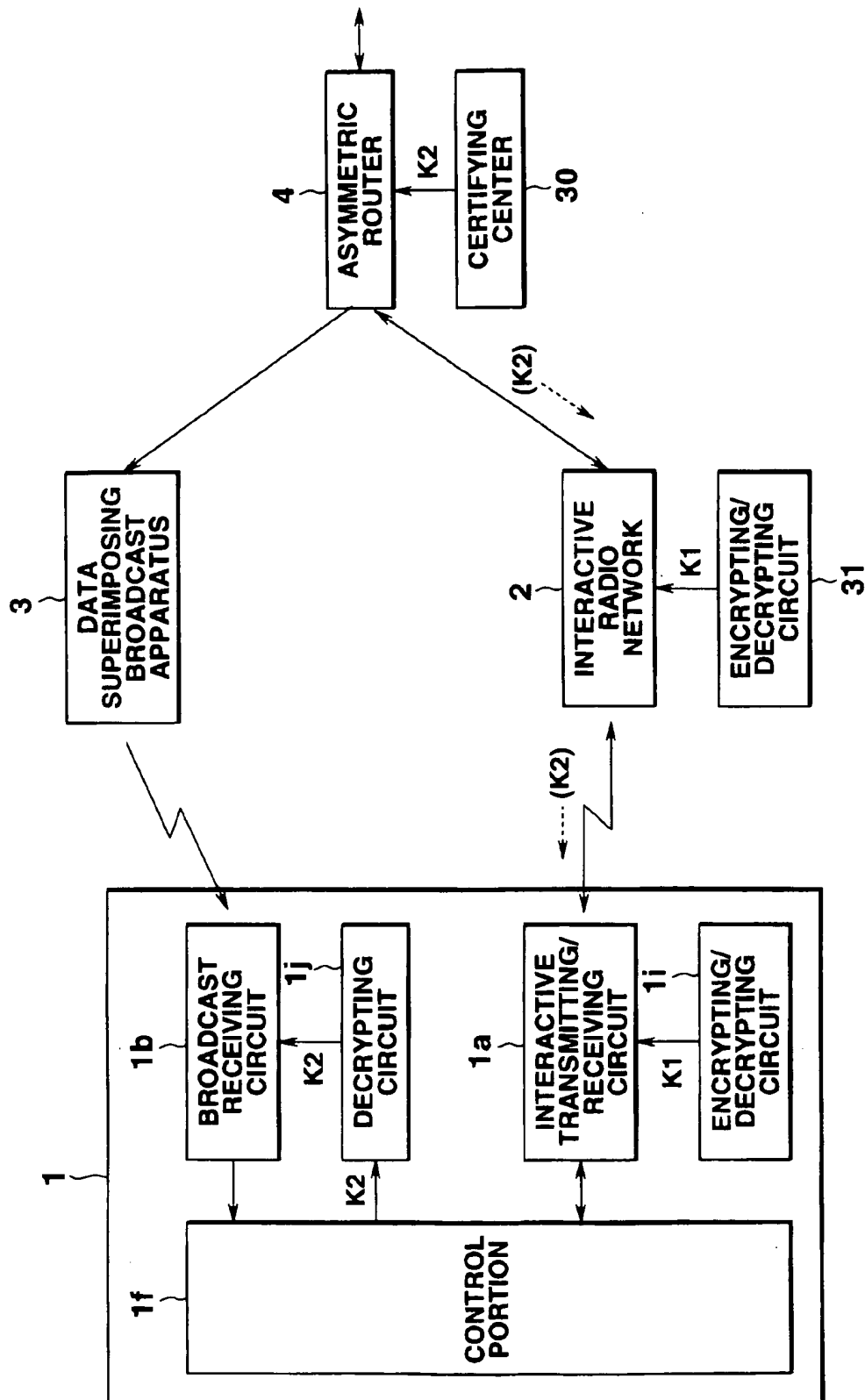


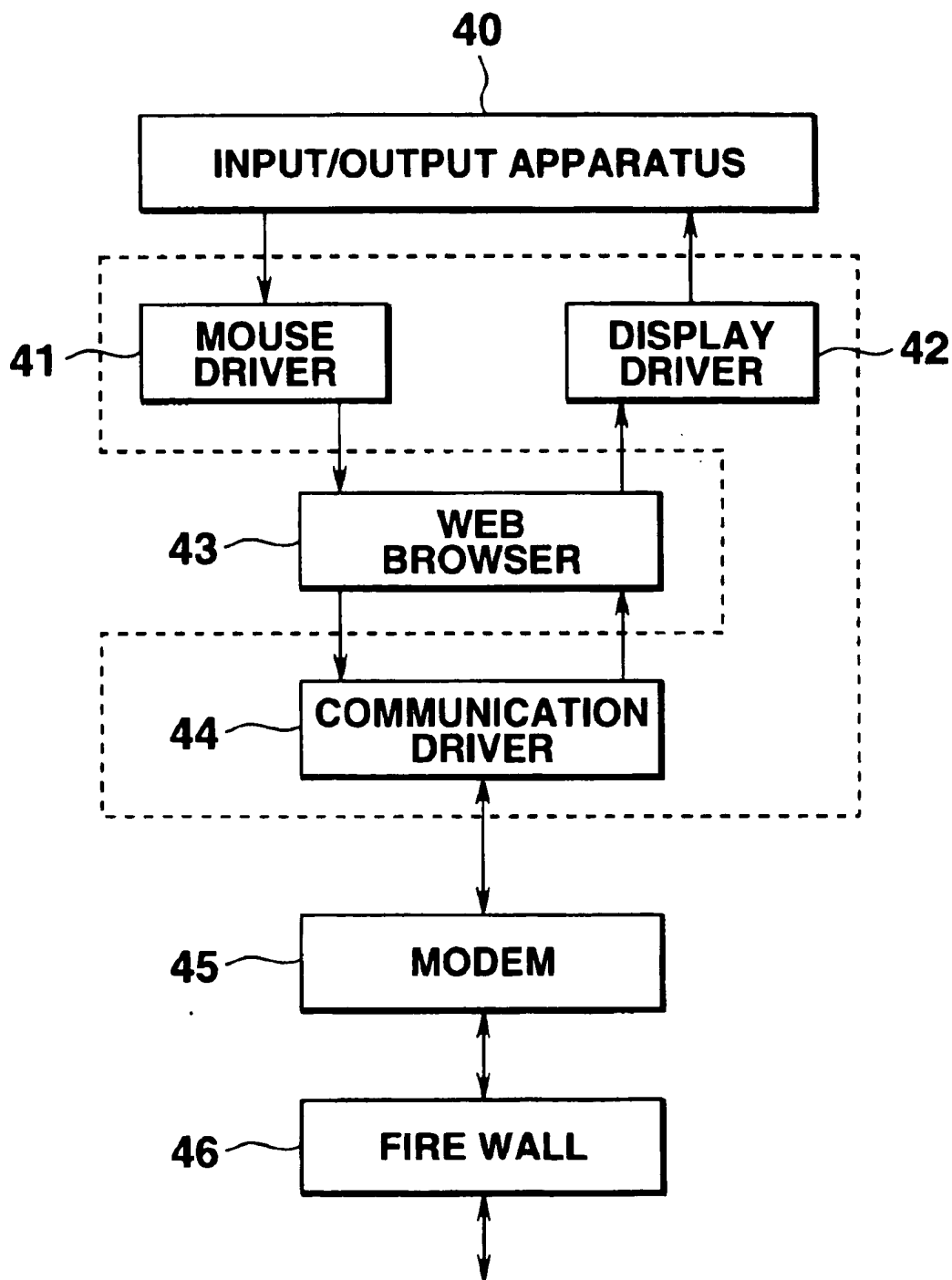
FIG.11

FIG. 12

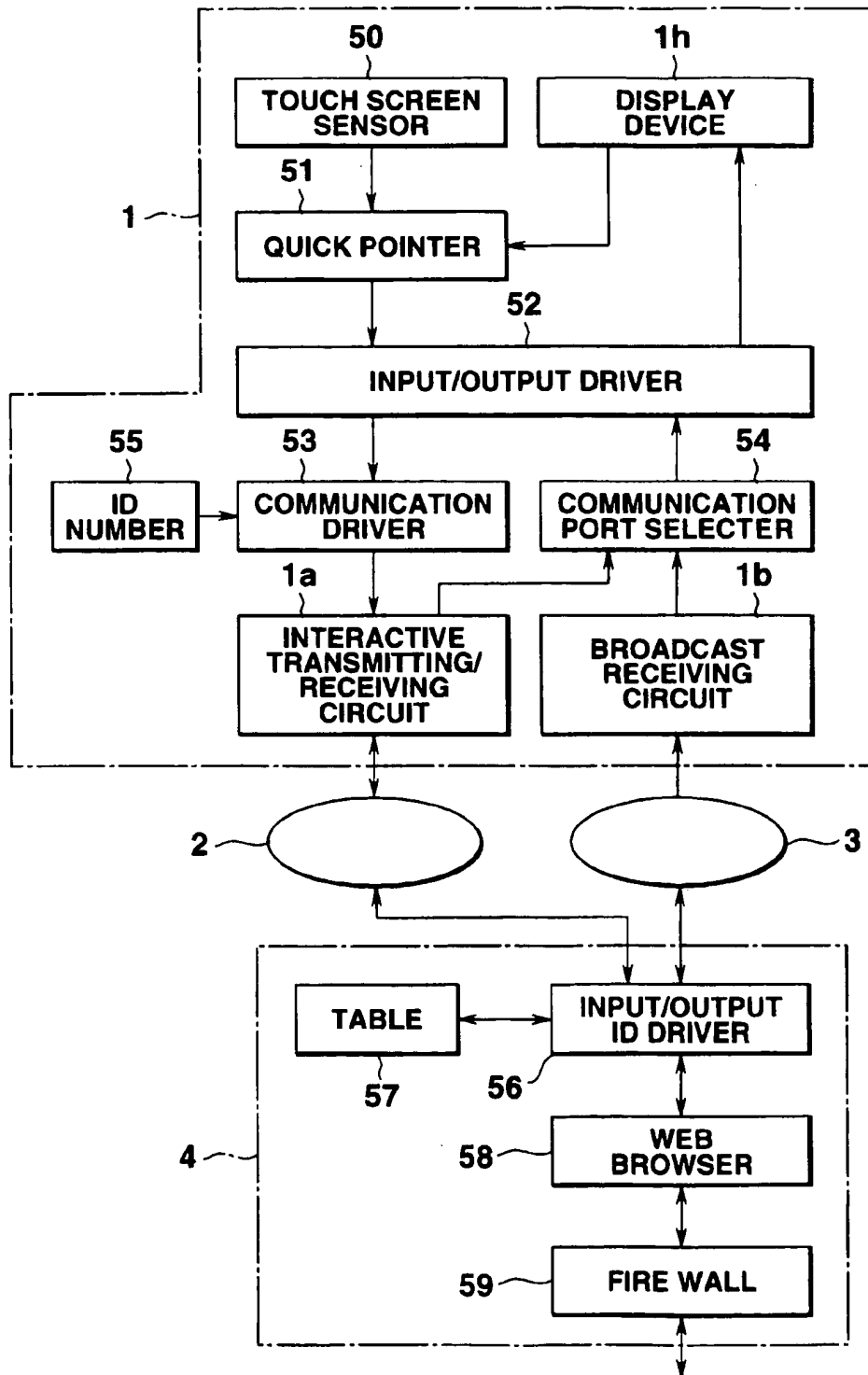


FIG. 13

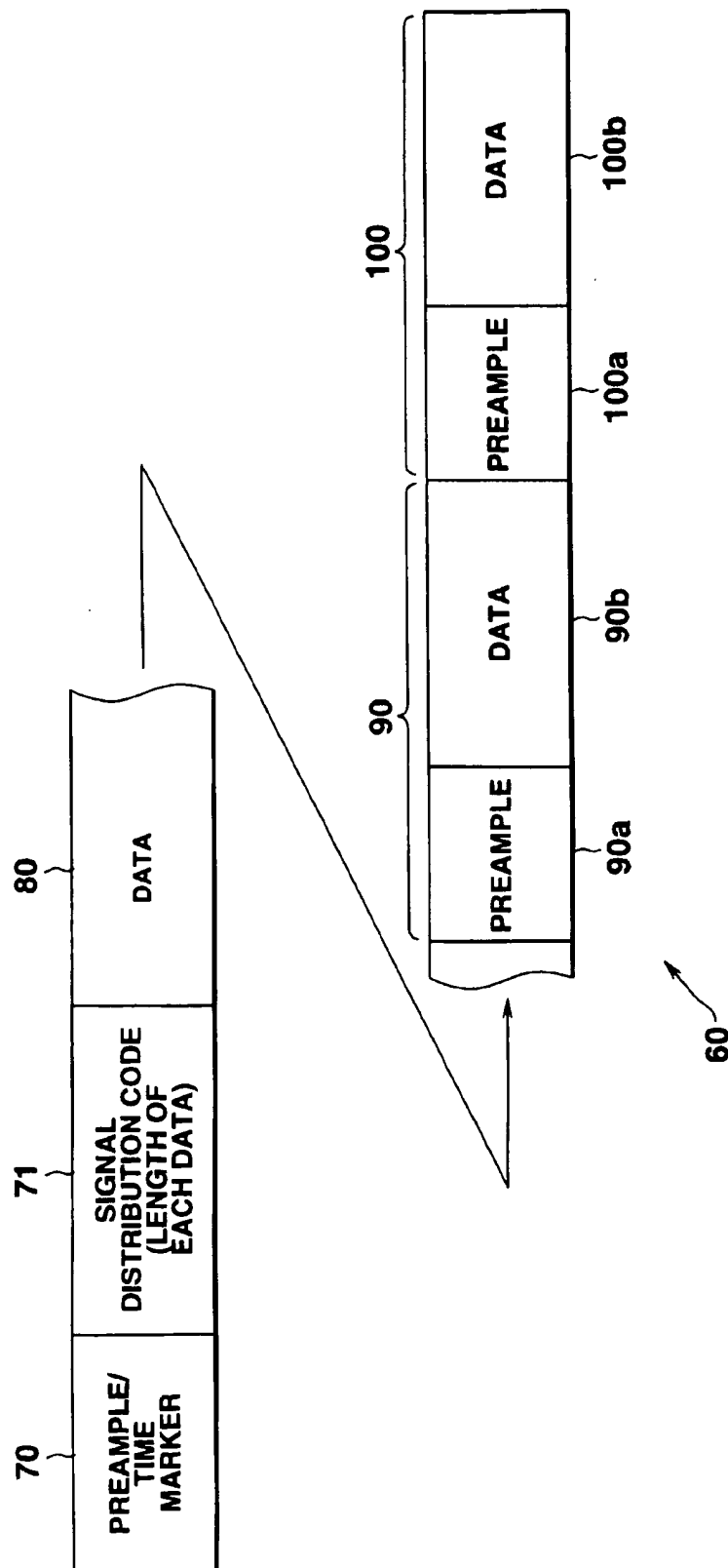
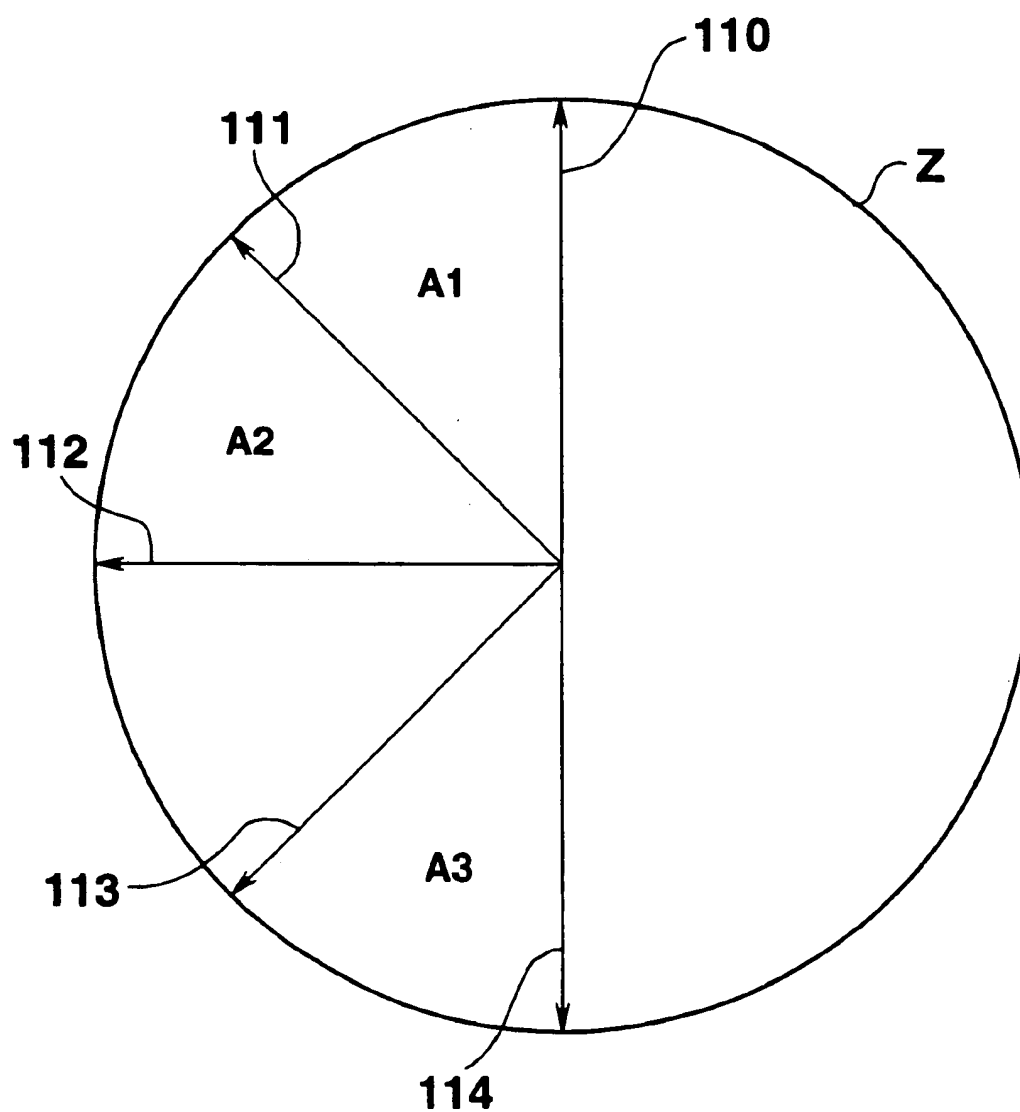


FIG.14

1

INTERACTIVE COMMUNICATION SYSTEM FOR DOWNLOADING LARGE AMOUNT DATA

BACKGROUND OF THE INVENTION

The present invention relates to an interactive communication system which transmits/receives data at a high speed using an existing communication system.

This application is based on Japanese Patent Application No.08-122057, filed May 16, 1996, the contents of which are incorporated herein by reference.

Recently, the internet (which uses a general analog telephone line) or a digital private line has become widely known as an interactive communication system. Using the internet, HTML (i.e., Hypertext Markup Language) formatted data (a document comprising an image, a sound and a character or text) can be browsed, and a number of large amount data which are stored in servers distributed all over the world can be received.

For interactive communication, there exists a usual television broadcast system for unidirectionally transmitting images and sound to a user over a plurality of channels. In a television broadcast, an empty portion of a television signal is used so that a character information (character broadcast) is transmitted by a plurality of channels. The television broadcast thus provides a communication system which can transmit a large amount of information to the user.

Furthermore, as an interactive communication system for transmitting/receiving more information between a plurality of users, there is such an interactive communication system as VOD (Video On Demand) using a private communication line network in which an up link/down link is provided in a CATV (cable television) system so that the interactive communication can be carried out.

In the case of accessing the internet using a fixed terminal, for example, a computer and the like provided at home or in an office, data can be transmitted at a high speed. However, in the case of using a portable type information terminal (a mobile terminal) carried by the user, since the data is transmitted via a relay station (a base station of the interactive communication network), a maximum data transfer rate is about 9.6 kbps. For example, the portable information terminal is used so as to access a WWW (World Wide Web) server on the internet and to browse HTML formatted data (i.e., documents comprising image, sound and character data). In this case, when the data of 150 kbytes is received, it takes about 15 seconds, and there is a problem in that the data transfer rate is too slow and is not practical.

In Europe, in a ground wave digital broadcast which is put to practical use as a television broadcast, a mobile communication which can realize a transfer rate of about 1 Mbps forms a part of the service. In the case of using this communication system, in the above example, the data can be received in about 2 seconds. However, in the television broadcast and the ground digital broadcast, even if the data can be transmitted at the high speed, the user can only receive the data. Thus, there is such a problem in that interactive communication cannot be carried out.

Furthermore, in such an interactive communication system as CATV using a private line, service cannot be provided to the user at an optional place (i.e., such as in a train, a hall or a conference center). In other words, CATV can be provided to the user only at a specific place. Thus, there is a problem in that thus service cannot be provided for a mobile type information terminal.

2

BRIEF SUMMARY OF THE INVENTION

Therefore, it is an object of the present invention to provide an interactive communication system which can transmit large amount data at a high speed to an information terminal, and more specifically, to a portable information terminal.

According to a first aspect of the present invention, there is provided an interactive communication system comprising:

- an interactive communication system in which data is transmitted/received between a host terminal and a terminal;
- a large amount data transmitting system for unidirectionally transmitting a large amount data to the host terminal; and
- a system switching apparatus for supplying data transmitted from the interactive communication system to the large amount data transmitting system so as to transmit the data to the host terminal via the large amount data transmitting system.

According to a second aspect of the present invention, there is provided an interactive communication system as recited in the first aspect in which the large amount data transmitting system comprises a broadcasting system for broadcasting a predetermined television signal, and when data is transmitted from the system switching apparatus, the data is superimposed to the television signal so that the data is broadcasted.

According to a third aspect of the present invention, there is provided an interactive communication system as recited in the first aspect in which the system switching apparatus, based on a data amount to be transmitted, selectively switches either the interactive communication system or the large amount data transmitting system for using as a data transmission channel.

According to a fourth aspect of the present invention, there is provided an interactive communication system as recited in the first aspect in which the large amount data transmitting system comprising:

- storage means for temporarily storing transmitted data;
- transmitting-scheduled time determining means for determining a scheduled time of broadcasting based on at least the data amount to be transmitted; and
- transmitting means for transmitting the data stored in the storage means when the scheduled time of broadcasting determined by the transmitting-scheduled time determining means comes.

According to a fifth aspect of the present invention, there is provided an interactive communication system as recited in the fourth aspect in which the large amount data transmitting system supplies the scheduled time of broadcasting determined by the transmitting-scheduled time determining means to the host terminal via the interactive communication system; and

the host terminal receives the scheduled time of broadcasting supplied from the interactive communication system, and

when the scheduled time of broadcasting comes, the data from the large amount data transmitting system is received.

According to a sixth aspect of the present invention, there is provided an interactive communication system as recited in the first aspect in which the large amount data transmitting system comprises plural data transmitting devices which are located in each predetermined area; and

the system switching apparatus temporarily stores supplied data, stored data being transmitted to one of the data transmitting devices located in an area corresponding to a position information of the host terminal.

According to a seventh aspect of the present invention, there is provided an interactive communication system as recited in the first aspect in which the system switching apparatus performs a predetermined encryption using a predetermined encrypting key relative to data to be transmitted to the large amount data transmitting system; and the host terminal decrypts encrypted large amount data transmitted from the large amount data transmitting system by using the predetermined encrypting key.

According to an eighth aspect of the present invention, there is provided an interactive communication system as recited in the seventh aspect in which the host terminal obtains the predetermined encryption key via the interactive communication system prior to a reception of the large amount data.

According to a ninth aspect of the present invention, there is provided an interactive communication system as recited in the first aspect in which the interactive communication system and the host terminal perform a predetermined encryption using a predetermined encrypting key relative to data to be transmitted/received, and encrypted data is decrypted by using the predetermined encrypting key.

According to a tenth aspect of the present invention, there is provided an interactive communication system as recited in the first aspect in which the host terminal comprises a plurality of server terminals for providing an HTML formatted file; and

the system switching apparatus browses the HTML formatted file provided by a predetermined server terminal on internet according to a command supplied from the server terminal via the interactive communication system, the HTML formatted file being converted into a display format data suitable for the server terminal so that the display format data is transmitted to the large amount data transmitting system.

According to an eleventh aspect of the present invention, there is provided an interactive communication system as recited in the sixth aspect in which the large amount data transmitting system transmits a communication information indicative of the interactive communication system which can be used in each area by the host terminal; and

the host terminal selects the interactive communication system to be used according to the communication information from the large amount data transmitting system.

According to a twelfth aspect of the present invention, there is provided an interactive communication system as recited in the first aspect in which the large amount data transmitting system transmits an identification code for specifying the host terminal by which the data is to be received for the data supplied from the system switching apparatus so as to transmit the data; and

the terminal determines whether or not the identification code provided for the data supplied from the large amount data transmitting system corresponds to the identification code which is preset to the identification code itself, the data being received only if a correspondence is detected.

According to a thirteenth aspect of the present invention, there is provided an interactive communication system as recited in the first aspect in which the host terminal supplies a position information indicative of a present position to the interactive communication system at a predetermined timing;

the system switching apparatus transmits the position information of the information terminal provided via the interactive communication system to the large amount data transmitting system; and

the large amount data transmitting system comprises a plurality of transmitting antennas having a different directivity, the data being transmitted from data transmission antenna having the corresponding directivity to the position information among the plurality of transmitting antennas.

Additional objects and advantages of the present invention will be set forth in the description which follows, and in part will be obvious from the description, or may be learned by practice of the present invention.

The objects and advantages of the present invention may be realized and obtained by means of the instrumentalities and combinations particularly pointed out in the appended claims.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING

The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate presently preferred embodiments of the present invention and, together with the general description given above and the detailed description of the preferred embodiments given below, serve to explain the principles of the present invention in which:

FIG. 1 is a block diagram showing a constitution of an interactive communication system according to a first embodiment of the present invention;

FIG. 2 is a block diagram showing a general constitution of a data superimposing broadcast apparatus according to the first embodiment of the present invention;

FIG. 3 is a block diagram showing a general constitution of an information terminal;

FIG. 4 schematically shows a layer structure of the information terminal;

FIG. 5 is a flow chart showing a part of a basic operation of an asymmetric router;

FIG. 6 is a flow chart showing a part of a basic operation of the data superimposing broadcast apparatus;

FIG. 7 is a flow chart showing a part of a basic operation of the information terminal;

FIG. 8 is a flow chart showing the operation of the interactive communication system according to the first embodiment;

FIG. 9 is a block diagram showing a constitution of an interactive communication system according to a second embodiment of the present invention;

FIG. 10 is a block diagram showing a constitution for realizing an encryption processing so as to transmit the data to a specific information terminal alone;

FIG. 11 is a block diagram showing a constitution of the information terminal comprising a conventional personal computer and the like;

FIG. 12 is a block diagram showing a general constitution of the information terminal and the asymmetric router according to a fourth embodiment;

FIG. 13 schematically shows a data format broadcasted by the data superimposing broadcast apparatus according to a sixth embodiment of the present invention; and

FIG. 14 schematically shows the operation in case of using a directional antenna according to a seventh embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

A preferred embodiment of an interactive communication system according to the present invention will now be described with reference to the accompanying drawings.

Embodiments according to the present invention will be described below as an embodiment applied to internet and a value added network (VAN).

(A) First Embodiment

(A-1) Constitution of an Interactive Communication System

FIG. 1 is a block diagram showing a constitution of an interactive communication system according to a first embodiment of the present invention. In FIG. 1, an information terminal 1 is carried by a user (not shown). The information terminal 1 comprises a key input device (not shown) for inputting an instruction from the user, a display device for displaying various data (image, character and the like) and the like. Furthermore, the information terminal 1 comprises a transmission/reception function for transmitting/receiving the data between the transmission/reception function and an interactive radio network 2 described below and a broadcast reception function for receiving the data from a data superimposing broadcast apparatus 3. The information terminal 1 carries out a usual interactive communication with the interactive radio network 2, whereby the information terminal 1 transmits/receives, via the interactive radio network 2, the data (small amount) such as a command, to/from a server 6 on an internet 5 and a value added network 8 on a fixed network 7 described above. Furthermore, the information terminal 1 receives the data (large amount) from the server 6 on internet 5 and the value added network 8 on the fixed network 7 via the data superimposing broadcast apparatus 3.

The interactive radio network 2 is a radio network which can perform the interactive communication with the information terminal 1 by radio. For example, the interactive radio network 2 is a communication network formed of base stations provided for each area which can be covered with a radio signal of a predetermined output level such as the communication network for a portable telephone, a PHS and the like. The interactive radio network 2 transmits the small amount data received from the information terminal 1 by radio to an asymmetric router 4 described later. Furthermore, the interactive radio network 2 transmits the small amount data from the asymmetric router 4 to the information terminal 1 by radio.

The data superimposing broadcast apparatus 3 is only unidirectional (transmitting). The data superimposing broadcast apparatus 3 is an apparatus for superimposing a predetermined large amount data from the asymmetric router 4 to a usual broadcasting signal and for transmitting the broadcasting signal. For example, the data superimposing broadcast apparatus 3 is a television broadcasting station for superimposing the large amount data provided from the asymmetric router 4 to the television signal such as a character broadcast and for broadcasting the television signal from a broadcasting tower or antenna (not shown). The data superimposing broadcast apparatus 3 may be a usual analog television broadcasting station or may be a digital television broadcasting station adopted in Europe as described above.

The asymmetric router 4 transmits the small amount data transmitted from the interactive radio network 2 to a predetermined server 6 on internet (a private line or an analog line) 5 described below and to the value added communication network 8 via the fixed network (the private line or the analog line) 7. The asymmetric router 4 transmits the

large amount data transmitted from the server 6 on internet 5 and the value added communication network on the fixed network 7 to the data superimposing broadcast apparatus 3.

Internet 5 is the communication network in which computers (servers) of various organizations and individuals all over the world are connected to one another. Internet 5 is directly connected to the asymmetric router 4 or is connected to the asymmetric router 4 through the private line or the usual analog line via a provider. The server 6 is a WWW (World Wide Web) server (hereinafter referred to as a Web server) for providing an HTML formatted data (a document comprising the image, a sound and the character) and the server for providing various news and data. Furthermore, the fixed network 7 is the private line or the analog line connected to the value added network 8 for providing various information, and it is directly connected to the asymmetric router 4.

(A-2) Data Superimposing Broadcast Apparatus

FIG. 2 is a block diagram showing a general constitution of the data superimposing broadcast apparatus 3. The data superimposing broadcast apparatus 3 comprises a memory 3a for temporarily storing the large amount data transmitted from the asymmetric router 4 and a superimposing apparatus 3b for superimposing the large amount data stored in the memory 3a to the television signal at a predetermined timing (scheduled time of broadcasting) and for transmitting the television signal. If the asymmetric router 4 issues a data transmission request, the data superimposing broadcast apparatus 3 supplies data indicating an idle channel to the asymmetric router 4, temporally stores the received large amount data in the memory 3a, and determines the scheduled time of broadcasting for transmitting the large amount data to the information terminal 1 according to a data amount and the transmission request. Then, the data indicating scheduled time of broadcasting is returned to the asymmetric router 4. At the scheduled time of broadcasting, the large amount data stored in the memory 3a is superimposed to the television signal by the superimposing apparatus 3b and is transmitted.

In order to determine the scheduled time of broadcasting, for example, a FIFO system for transmitting the data in an order of the previously received data and such a system that a priority order is combined to this order so as to determine a transmission order are preferable. Furthermore, when a simulcast is previously scheduled, the scheduled time of broadcasting may be determined to avoid the simulcast timing. Furthermore, the data stored in the memory 3a is prepared for a re-transmission, and the data is maintained even after the transmission. The data is erased when an ACK signal transmitted from the information terminal 1 is received via the asymmetric router 4 or a predetermined time passes.

(A-3) Information Terminal

FIG. 3 is a block diagram showing a general constitution of the information terminal. The information terminal 1 has an interactive transmitting/receiving circuit 1a for transmitting/receiving the data relative to the interactive radio network 2 and a broadcast receiving circuit 1b for receiving the television signal from the data superimposing broadcast apparatus 3. According to the scheduled time of broadcasting which is previously transmitted via the interactive radio network 2, the information terminal 1 is operated in such a manner that a reception preparation and a data selection are carried out. The broadcast receiving circuit 1b is in a sleep state in order to reduce power consumption at a time zone except for a receiving operation. Based on the time measured by a timer 1e, the broadcast receiving circuit

1b is started by a control circuit 1f. The broadcast receiving circuit 1b is operated so as to receive the television signal to which the large amount data is superimposed. Furthermore, this starting control may include a function of specifying a reception frequency (channel).

A data selecting circuit 1c selects the data received by either the interactive radio network 2 or the data superimposing broadcast apparatus 3 under the control of the control circuit 1f, and it transmits the data to a processing apparatus 1d. The processing apparatus 1d processes the supplied data, and the data is converted to a predetermined display format data, so that the data is displayed in a display device 1h.

Based on a system clock (not shown), the timer 1e measures a present time, and it transmits data indicating the present time to the control circuit 1f. The control circuit 1f controls each of the above circuits/devices. Furthermore, the control circuit 1f compares the present time with the scheduled time of broadcasting which is received via the interactive radio network 2. When the scheduled time of broadcasting comes, the broadcast receiving circuit 1b starts operation so that the television signal to which the large amount data transmitted from the data superimposing broadcast apparatus 3 is superimposed is received.

Furthermore, a key input device 1g includes a keyboard for accessing internet 5, for accessing the value added network 8 and for giving various instructions to the control circuit 1f. Furthermore, the display device 1h comprises, for example, a liquid crystal display device. The display device 1h displays a home page provided on internet 5 and the information of the value added network 8. As the display device 1h, such a display device as a CRT may be externally connected.

(A-4) Layer Structure of the Information Terminal

FIG. 4 is a schematic diagram showing a layer structure of the information terminal 1. The information terminal 1 comprises a physical layer 10 relative to the interactive radio network 2, a data link layer 11, a network layer 12, further, a physical layer 13 relative to the data superimposing broadcast apparatus 3, a data link layer 14. The information terminal 1 further comprises an asymmetric routing layer 15 which is over the network layer 12 and the data link layer 14, and an application layer 16 which is over the asymmetric routing layer 15. The application layer 16 is a Web browser (software), for example, in case of accessing internet 5. Since a routing (switching) between the interactive radio network 2 and the data superimposing broadcast apparatus 3 is carried out by the asymmetric routing layer 15, when a receiver is to be selected, it is not necessary for the application layer 16 to recognize a plurality of receivers.

(B) Operation of the First Embodiment

A part of a basic operation of the asymmetric router 4, the data superimposing broadcast apparatus 3 and the information terminal 1 will be described. FIGS. 5 to 7 are flow charts showing a part of the basic operation of the asymmetric router, the data superimposing broadcast apparatus 3 and the information terminal 1, respectively.

(B-1) Asymmetric Router

A procedure of selecting a channel of a downlink in the asymmetric router 4 will be described with reference to FIG. 5. When the asymmetric router 4 receives the data (large amount or small amount), at step S10, the amount of the received data is calculated. It is determined, at step S12, whether or not the total amount of the data is larger than a preset value. The preset value is a threshold value for determining whether the data to be transmitted to the information terminal 1 should be transmitted by the data superimposing broadcast apparatus 3 or by the interactive radio

network 2. When the data has the large amount, it is more effective to be transmitted via the data superimposing broadcast apparatus 3 having a relatively high-speed data transfer rate. When the data has the small amount, it is more effective to be transmitted via the interactive radio network 2 having a relatively low-speed data transfer rate. Therefore, when the total amount of the data is the preset value or less, the process proceeds to step S14. At step S14, the data is transmitted to the information terminal 1 via the interactive radio network 2.

When the total amount of the data is larger than the preset value, the process proceeds to step S16. At step S16, the transmission request is transmitted to the data superimposing broadcast apparatus 3. At step S18, it is determined whether or not the transmission is possible based on a response from the data superimposing broadcast apparatus 3. For example, the transmission request specifies the channel by which the data should be transmitted. When the transmission is possible, data indicating the channel by which the data should be transmitted is returned. When the transmission is impossible, the process proceeds to step S14. The data is transmitted to the information terminal 1 via the interactive radio network 2.

When the transmission is possible, the process proceeds to step S20. At step S20, the data is transmitted via the data superimposing broadcast apparatus 3. When the data superimposing broadcast apparatus 3 receives the data, the scheduled time of broadcasting is determined based on a data amount, and the data is returned to the asymmetric router 4. At step S22, the asymmetric router 4 receives the scheduled time of broadcasting. At step S24, data transmission channel and the scheduled time of broadcasting are supplied to the information terminal 1 via the interactive radio network 2.

At step S26, it is waited that the ACK signal indicative of a completion of a data reception is supplied from the information terminal 1 via the interactive radio network 2. When the ACK signal is received, the received ACK signal is transmitted to the data superimposing broadcast apparatus 3 at step S28 and then the process is completed.

(B-2) Data Superimposing Broadcast Apparatus

A broadcasting procedure of the downlink in the data superimposing broadcast apparatus will be described with reference to FIG. 6. When the data superimposing broadcast apparatus 3 receives the data from the asymmetric router 4, the data is temporally stored in the memory 3a at step S30. At step S32, the received data amount is calculated. At step S34, the scheduled time of broadcasting is determined based on the obtained data amount. The scheduled time of broadcasting is transmitted to the asymmetric router 4.

At step S38, it is determined whether or not the scheduled time of broadcasting comes. If the scheduled time of broadcasting is not reached, step S38 is repeated, and it is in a waiting mode. When it is the scheduled time of broadcasting, the process proceeds to step S40. At step S40, the data stored in the memory 3a at step S30 is superimposed on an empty portion of the usual television signal and it is broadcasted. At step S42, it is determined whether or not the ACK signal from the asymmetric router 4 is received. The ACK signal is a confirmation signal indicative of receiving the whole data transmitted to the asymmetric router 4 by the information terminal 1 receiving the television signal to which the data is superimposed. The ACK signal is transmitted to the data superimposing broadcast apparatus 3 via the asymmetric router 4.

If the ACK signal from the asymmetric router 4 is not received, the process proceeds to step S44 in which it is determined whether or not it is time out. If not, the process

is returned to step S42 in order to wait until the ACK signal is received. When the ACK signal cannot be received after a superimposing broadcast of the data, it is time out. In this case, the process proceeds to step S46, and a not-received notification is transmitted to the asymmetric router 4 in order

After the superimposing broadcast of the data, when the ACK signal is received from the information terminal 1 via the asymmetric router 4 at step S42, it can be determined that the data is received by the information terminal 1. The process proceeds to step S48 in which the broadcasted data is erased from the memory 3a. Then, the process is completed.

(B-3) Information Terminal

The procedure of selecting a reception path in the information terminal will be described with reference to FIG. 7. At step S50, the information terminal 1 determines whether a scheduled channel is the interactive radio network 2 or the data superimposing broadcast apparatus 3. When the scheduled channel is the interactive radio network 2, the process proceeds to step S52. The interactive transmitting/receiving circuit (receiving circuit) 1a being a receiving function of the interactive radio network 2 is selected and it is started. At step S54, the small amount data is received from the interactive radio network 2.

When the scheduled channel is the data superimposing broadcast apparatus 3, the process proceeds to step S56. The broadcast receiving circuit 1b being the receiving function of the data superimposing broadcast apparatus 3 is selected. At step S58, the channel to be broadcasted notified from the asymmetric router 4 is set. At step S60, with reference to the present time of the timer 1e, it is determined whether or not the scheduled time of broadcasting notified from the asymmetric router 4 comes. If the scheduled time of broadcasting is not reached, step S60 is repeated, and it is in the waiting mode. When the scheduled time of broadcasting comes, the process proceeds to step S62. The sleep of the broadcast receiving circuit 1b is released. At step S64, the television signal from the data superimposing broadcast apparatus 3 to which the data is superimposed is received.

When the reception of the data superimposed to the television signal is completed, the broadcast receiving circuit 1b is returned to the sleep state at step S66. At step S68, after the ACK signal indicative of the completion of the data reception is transmitted to the asymmetric router 4 via the interactive radio network 2, the process is completed.

(B-4) Whole Operation

A whole operation of the first embodiment will be described. FIG. 8 is a sequence for explaining the operation of the interactive communication system according to the first embodiment. When the transmission request is transmitted from the Web browser to the Web server 6 on internet 5, the information terminal 1 transmits a command CMD to the asymmetric router 4 via the interactive radio network 2. The asymmetric router 4 transmits the command CMD to the Web server 6 via internet 5. According to the received command CMD, the Web server 6 transmits, for example, a browser data (an HTML formatted file) D1 to the asymmetric router 4 via internet 5.

When a data amount of the browser data D1 transmitted from the Web server 6 exceeds a predetermined amount, the asymmetric router 4 transmits a channel requirement CRQ to which the browser data D1 should be transmitted to the data superimposing broadcast apparatus 3. The data superimposing broadcast apparatus 3 returns data indicating the idle channel and the ACK signal to the asymmetric router 4.

When the asymmetric router 4 receives the ACK signal, the browser data D1 is transmitted to the data superimposing broadcast apparatus 3. When the data superimposing broadcast apparatus 3 receives the browser data D1, the scheduled time of broadcasting is determined in accordance with the data amount. The scheduled time of broadcasting is transmitted to the asymmetric router 4 as a time tag TT.

When the asymmetric router 4 receives the time tag TT, data indicating a previously received transmitting channel C and the time tag TT are transmitted to the information terminal 1 via the interactive radio network 2. The information terminal 1 receives the data indicating data transmission channel C and the time tag TT by the interactive transmitting/receiving circuit 1a. According to the time tag TT, the information terminal 1 waits until the television signal to which the browser data D1 is superimposed is transmitted.

When the scheduled time of broadcasting comes, the data superimposing broadcast apparatus 3 superimposes the browser data D1 to the television signal of the preset transmitting channel C by using the superimposing apparatus 3b, so that the television signal is transmitted. Similarly, when the scheduled time of broadcasting comes, the information terminal 1 starts the broadcast receiving circuit 1b. The television signal to which the browser data D1 is superimposed is received by the started broadcast receiving circuit 1b. The browser data D1 is extracted from the television signal. A predetermined formatted display data is generated. The data is displayed in the display device 1h. Furthermore, the information terminal 1 transmits the ACK signal indicative of the reception completion of the browser data D1 to the interactive transmitting/receiving circuit 1a. The ACK signal is transmitted to the asymmetric router 4 via the interactive radio network 2 by the interactive transmitting/receiving circuit 1a. When the asymmetric router 4 receives the ACK signal, the ACK signal is transmitted to the data superimposing broadcast apparatus 3. The data superimposing broadcast apparatus 3 erases the browser data D1 stored in the memory 3a.

In such a manner, according to the first embodiment, a data stream (the small amount data such as the command) in an upstream direction relative to the information terminal 1 is transmitted via the interactive radio network 2 whose data transfer rate is relatively low. On the other hand, a data stream (the large amount data such as the browser data) in a downstream direction relative to the information terminal 1 is transmitted by the data superimposing broadcast apparatus 3 whose data transfer rate is relatively high. Therefore, a load of the relatively low-speed interactive radio network 2 is reduced, and further the data can be transmitted at a high speed.

(C) Second Embodiment

A second embodiment of the present invention will be described. According to the second embodiment, among a plurality of data superimposing broadcast apparatuses, the data superimposing broadcast apparatus in which a present position of the information terminal 1 is defined as a service area (radio wave arriving area) is selected. The television signal to which the large amount data is superimposed is transmitted from the selected data superimposing broadcast apparatus. Furthermore, when the data amount to be transmitted exceeds the amount which the data superimposing broadcast apparatus 3 can broadcast, the large amount data is transmitted via the interactive radio network 2. The constitution and the operation of the second embodiment will be described below.

FIG. 9 is a block diagram showing the constitution of the interactive communication system according to the second

embodiment of the present invention. The same portions as those of the first embodiment will be indicated in the same reference numerals and their detailed description will be omitted. An asymmetric router 20 comprises a memory 20a for temporally storing the large amount data received via internet 5 or the fixed network 7. The large amount data is usually provided for any one of data superimposing broadcast apparatuses 3-1, 3-2, . . . , 3-n. The data is transmitted to the information terminal 1 by the data superimposing broadcast apparatus 3-i (i=1 to n). However, when the data amount to be transmitted exceeds the amount which the data superimposing broadcast apparatus 3-i can broadcast, the large amount data must be transmitted via the interactive radio network 2. In this case, since the interactive radio network 2 has a relatively low data transfer rate, the large amount data is needed to be temporally stored. As described above, when the data amount cannot be transmitted by the data superimposing broadcast apparatus 3-i, the memory 20a of the asymmetric router 20 is used for transmitting via the interactive radio network 2.

Furthermore, each of the data superimposing broadcast apparatuses 3-1 to 3-n is provided in each region to which a predetermined outputted broadcast wave can arrive. The large amount data is superimposed to the television signal and broadcasted by the apparatus in which the present position of the information terminal 1 is defined as the service area. The asymmetric router 20 obtains the present position of the information terminal 1 via the base stations provided in each predetermined area in the interactive radio network 2. The asymmetric router 20 selects based on the obtained present position of the information terminal 1 to which the data superimposing broadcast apparatuses 3-1 to 3-n the large amount data is supplied. The asymmetric router 20 is previously provided with a correspondence as a table between the present position of the information terminal 1, that is, the area where the information terminal 1 exists which is obtained via the base station and the data superimposing broadcast apparatus 3 in which the area is defined as the service area.

(D) Operation of the Second Embodiment

When the asymmetric router 20 receives the command for a connection requirement to the Web server 6 on internet 5 from the information terminal 1 via the interactive radio network 2, the command is transmitted to the Web server 6 via internet 5. When the command is received, a position information indicative of the present position of the information terminal 1 is obtained. When the data (browser data) of large amount transmitted from the Web server 6 on internet 5 is received, the asymmetric router 20 temporarily stores the received browser data in the memory 20a.

Based on the position information of the information terminal 1, the asymmetric router 20 requires for the channel by which the browser data is transmitted to the corresponding data superimposing broadcast apparatus 3-i among the data superimposing broadcast apparatuses 3-i to 3-n. When data transmission to the corresponding data superimposing broadcast apparatus 3-i is possible, the browser data which is temporarily stored in the memory 20a is transmitted to the corresponding data superimposing broadcast apparatus 3-i in the same way as the first embodiment. Thereafter, similarly to the first embodiment, the data superimposing broadcast apparatus 3-i superimposes the browser data to the television signal and it broadcasts the signal at a predetermined scheduled time of broadcasting.

When data transmission to the corresponding data superimposing broadcast apparatus 3-i is impossible, the asymmetric router 20 notifies the information terminal 1 that the

interactive radio network 2 should be used to transmit the browser data. Thereafter, the browser data which is temporarily stored in the memory 20a is transmitted to the information terminal 1 via the interactive radio network 2. According to the notification from the asymmetric router 20, as shown in the flow charts of FIGS. 5-7, the information terminal 1 sets the reception path in the interactive radio network 2. The interactive transmitting/receiving circuit 1a receives the browser data. A display data is generated from the browser data. The display data is displayed in the display device 1h. When the asymmetric router 20 receives the ACK signal indicative of the reception completion of the browser data from the information terminal 1, the browser data in the memory 20a is erased.

As described above, according to the second embodiment, the large amount data is superimposed to the television signal and broadcasted (transmitted) by the data superimposing broadcast apparatus 3-i alone in which the present position of the information terminal 1 is defined as the service area among the plural data superimposing broadcast apparatuses 3-i to 3-n. Therefore, a resource of the radio wave can be effectively used. Furthermore, if there is no data superimposing broadcast apparatus which can transmit the data, the large amount data is temporarily stored in the memory 20a provided in the asymmetric router 20. The data stored in the memory 20a is transmitted to the information terminal 1 via the interactive radio network 2. Therefore, even if the data superimposing broadcast apparatus is occupied, it is not necessary to wait for a long time, and the large amount data can be received.

(E) Third Embodiment

A third embodiment of the present invention will be described. According to the third embodiment, in order that the data (the small amount and the large amount) may be transmitted/received between specific information terminals alone, the data is encrypted so as to be transmitted/received.

FIG. 10 is a block diagram showing a constitution for realizing an encryption processing so as to transmit the data to the specific information terminal alone. The same portions as those of the first embodiment will be indicated in the same reference numerals and their detailed description will be omitted. A certifying center 30 has an inherent encrypting key K for the asymmetric router 4, and the encrypting key K is supplied to the asymmetric router 4. The asymmetric router 4 encrypts the large amount data to the data superimposing broadcast apparatus 3 by using the encrypting key K. Based on the request from the information terminal 1, the asymmetric router 4 transmits the encrypting key K to the information terminal 1 via the interactive radio network 2.

An encrypting/decrypting portion 31 includes an inherent encrypting key K1 for the interactive radio network 2. The encrypting key K1 is supplied to the interactive radio network 2. The interactive radio network 2 uses the encrypting key K1, so that the small amount data between the information terminals is encrypted and decrypted.

The information terminal 1 comprises an encrypting/decrypting portion 1i for encrypting/decrypting the small amount data to be transmitted/received between the information terminal 1 and the interactive radio network 2 by using the encrypting key K1. Furthermore, the information terminal 1 comprises a decrypting portion 1j for decrypting the large amount data from the data superimposing broadcast apparatus 3 by using an encrypting key K2 obtained from the asymmetric router 4 via the interactive radio network 2.

It is assumed that the interactive radio network 2 encrypted by the encrypting key K1 being a safe path can be

used. The certifying center 30, preferably, uses "One Time Pad encryption system" using the time tag and the like, as the encryption system. The "One Time Pad encryption system" always generates the encrypting key for encrypting/decrypting. Furthermore, the encrypting key is used only once, and the same encrypting key is not used repeatedly. That is, whenever the data necessary for the encryption is generated, the encrypting key K2 is generated. The encrypting key K2 is used so that the data superimposing broadcast is encrypted. Thereafter, the encrypting key K2 is transmitted to the information terminal 1 via the interactive radio network 2, and the same-valued encrypting key K2 is never used.

According to the third embodiment, the data to be transmitted/received is encrypted between the asymmetric router 4 and the data superimposing broadcast apparatus 3, between the interactive radio network 2 and the information terminal 1 and between the data superimposing broadcast apparatus 3 and the information terminal 1, whereby the data can be safely transmitted. Furthermore, since the data can be received by the information terminal 1 which is previously authorized by the certifying center 30, for example, a charge management for each information terminal can be realized, which is very effective for providing a service.

(F) Fourth Embodiment

A fourth embodiment of the present invention will be described. According to the fourth embodiment, the constitution for reducing the load of the information terminal is provided. In general, for example, in case of accessing the Web server 6 on internet 5, in a general personal computer (or the terminal for accessing internet), for example, the Web browser is implemented on an OS environment such as Windows (trademark), whereby the access can be realized. However, in this case, since a sufficient memory amount and a high-speed processing are needed, a cost of the terminal is increased. Furthermore, a circuit constitution becomes complicated and large-scaled, so that portability is lost.

FIG. 11 is a block diagram showing the constitution using the terminal such as a conventional personal computer. An input/output apparatus 40 is a display (a liquid crystal display and the like) and a mouse/keyboard (which may be a cursor key and a track ball provided in a housing). A mouse driver 41 and a display driver 42 on the OS control the input/output apparatus 40. A Web browser 43 is application software operated under the OS. The Web browser 43 generates the predetermined format display data from the browser data. The Web browser 43 displays the data in the display device by using the display driver 42. The Web browser 43 receives the input signal from the mouse/keyboard operated by the user.

Furthermore, at the access to internet 5, the Web browser 43 transmits a specification (command) of a browsing page, the instruction (command) to download a program assigned to the page and the like to a communication driver 44. The communication driver 44 is software on the OS. The communication driver 44 controls a modem 45 (or an Ethernet card and the like), and it transmits the command to internet 5. In some cases, the connection to internet 5 is carried out via a fire wall 46 such as a Proxy server. The fire wall 46 is a security software for preventing an external invasion (access) to the information terminal 1. According to the information terminal comprising the conventional constitution, in addition to the OS, the software of large amount such as the Web browser must be started. Therefore, the memory amount and a processing rate are limited, so that it is difficult to realize a comfortable environment.

According to the fourth embodiment, the asymmetric router 4 carries out the processing of the Web browser, and

the information terminal 1 carries out only an input/output processing. The fourth embodiment will be described in detail below. FIG. 12 is a block diagram showing a general constitution of the information terminal and the asymmetric router according to the fourth embodiment. The same portions as those of the first embodiment (FIG. 3) will be indicated in the same reference numerals and their detailed description will be omitted.

The display device 1h is the display device comprising the liquid crystal display device and the like, as described above. The predetermined format display data (the large amount data: the browser data, the small amount data: the command) supplied via an input/output driver 52 is displayed. A touch screen sensor 50 is a sensor for sensing a touch on a screen of the display device 1h by the user. A touched position is supplied to a quick pointer 51. According to the touched position from the touch screen sensor 50 and a display screen information of the display device 1h, the quick pointer 51 obtains which position is touched on the display screen (position information). The position information is supplied to a communication driver 53 via the input/output driver 52.

The communication driver 53 supplies the position information as the command to the interactive transmitting/receiving circuit 1a. On demand, the communication driver 53 may supply an ID number 55 for identifying the information terminal 1 to the interactive transmitting/receiving circuit 1a. A communication port selector 54 selectively switches the small amount data supplied from the interactive radio network 2 via the interactive transmitting/receiving circuit 1a or the large amount data (browser data) from the data superimposing broadcast apparatus 3 received by the broadcast receiving circuit 1b. The data is supplied to the display device 1h via the input/output driver 52.

In the asymmetric router 4, an input/output ID driver 56 shown in FIG. 12 receives the data and the ID number transmitted from the information terminal 1 via the interactive radio network 2. Referring to a correspondence table 57 which stores a correspondence between the ID number and an address of internet, the data is transmitted to a Web browser 58. Furthermore, the large amount data (such as the browser data) received by the Web browser 58 is supplied to the data superimposing broadcast apparatus 3. As described above, the Web browser 58 is the software for accessing internet 5. According to the data from the input/output ID driver 56, the Web browser 58 accesses a predetermined server 6 on internet 5. Furthermore, the connection to internet 5 may be carried out via a fire wall 59 such as the Proxy server as usual.

According to the fourth embodiment, the asymmetric router 4 comprises the software for accessing internet 5. The information terminal 1 comprises the software alone for the input/output processing. Therefore, the load of the memory amount and the processing rate in the information terminal 1 can be reduced, so that the apparatus can be used in the comfortable environment. Thus, the user can access the server 6 on internet 5 by the inexpensive and light weight information terminal 1. Furthermore, by the information terminal which does not have the address of internet 5, the user can access the server 6 on internet 5.

(G) Fifth Embodiment

A fifth embodiment of the present invention will be described. According to the fifth embodiment, as the interactive radio network 2, a plurality of interactive radio networks such as the PHS which can provide the service in a predetermined area and a cellular telephone (portable telephone) which can provide the service almost all over the

area can be used. Based on the present position of the mobile information terminal 1, the available interactive radio network can be selected.

The fifth embodiment comprises the same as the second embodiment. A radio network identification data indicative of the available interactive radio network in the corresponding area is superimposed to each television signal transmitted from a plurality of data superimposing broadcast apparatuses provided in each area. When the information terminal 1 is moved to another area, it is confirmed the radio network identification data superimposed to the television signal received at a place to which the information terminal 1 is moved. Therefore, the available interactive radio network can be selected in the present area.

Furthermore, when the information terminal 1 is moved, a fact that the television signal can be received at the place to which the information terminal 1 is moved may be notified to the asymmetric router 4 via the interactive radio network 2. Therefore, the asymmetric router 4 can interactively transmit/receive the data relative to the information terminal 1 via the notified interactive radio network 2.

Furthermore, the above two systems may be combined to each other. The radio network identification data indicative of the available interactive radio network 2 in the corresponding area is superimposed to each television signal transmitted from a plurality of data superimposing broadcast apparatuses. A fact that the moved information terminal 1 can receive the television signal may be notified to the asymmetric router 4 via the interactive radio network 2 indicated by radio network identification data which is superimposed to the television signal and is broadcasted.

According to the fifth embodiment, even if the information terminal 1 is moved and it is moved to the service area for the different interactive radio network 2 and data superimposing broadcast apparatus 3, the appropriate interactive radio network 2 can be selected based on the present position of the information terminal 1. Therefore, it is possible for the information terminal 1 to move over a wide range.

(H) Sixth Embodiment

A sixth embodiment of the present invention will be described. According to the sixth embodiment, the data format broadcasted by the data superimposing broadcast apparatus 3 comprises the constitution shown in FIG. 13, so that a finer service is provided. In FIG. 13, a frame 60 comprises a preamble/time marker 70, a signal distribution code (length of each data) 71, a data 80 related to an individual user, a data 90 related to a group user and a public data 100.

The length of one frame 60 is variable so that the length of each data can be flexibly set. In order to reduce the power consumption, the frame 60 comprises a periodical time marker. According to the time marker, a flexible frame constitution is realized. The signal distribution code 71 specifies a distribution of each length of the data 80 related to the individual user, the data 90 related to the group user and the public data 100. The signal distribution code 71 is encoded.

The data 80 related to the individual user is encrypted so that a specific individual can decrypt the encrypted data by the key which is specified by an individual ID of the individual. A detail of the encryption is described in the third embodiment. The data 90 related to the group user is encrypted so that a common group can decrypt the encrypted data by the key specified by a group ID of the group. The data 90 related to the group user comprises a preamble 90a at the head thereof. The public data 100 is not encrypted so that many and unspecified receivers (information terminals)

can freely receive the data. Similarly to the data 90 related to the group user, the public data 100 comprises a preamble 100a at the head thereof.

The individual ID may be different from the group ID. The group ID may be embedded into a predetermined position of the individual ID. In an embedding form, for example, the ID form is defined as "AAAABB", the former "AAAA" is used for the individual ID for identifying the individual, and the latter "BB" is used for the group ID for identifying the group.

According to the sixth embodiment, the data which is superimposed to the television signal and transmitted from the data superimposing broadcast apparatus 3 is formed by the data 80 related to the individual user, the data 90 related to the group user and the public data 100, whereby the broadcast-authorized user can be flexibly set. The data 80 related to the individual user and the data 90 related to the group user are encrypted so that the broadcast-authorized user alone can decrypt the data, whereby a leakage of the information can be prevented. Furthermore, the time marker 70 is added, so that the information terminal 1 starts the reception by matching to a reception timing. Therefore, at the time except for the reception timing, the information terminal 1 is in the sleep state, whereby the power consumption can be reduced.

(I) Seventh Embodiment

A seventh embodiment of the present invention will be described. According to the seventh embodiment, in the data superimposing broadcast apparatus 3, a plurality of directional antennas directed in different directions are used, whereby radio waves of a common frequency can be used for each antennas, so that the data related to respective users are simultaneously transmitted to a plurality of users. As described in the sixth embodiment, in data transmission of the data related to the individual user, in order to communicate with the specific individual user via the interactive radio network 2, the service area corresponding to the base station number is known, so that the position of the individual user can be specified.

When the large amount data is simultaneously transmitted to a plurality of individual users, a plurality of directional antennas directed in different directions are used for the data superimposing broadcast apparatus 3. Therefore, if a plurality of individual users are sufficiently separated from one another in a direction which can sufficiently ensure a CIR (Carrier-to-Interference Ratio) by a directivity of the antenna, the different data can be simultaneously transmitted to the corresponding user by the common frequency.

FIG. 14 schematically shows the operation in case of using the directional antenna according to the seventh embodiment. In FIG. 14, a reference symbol Z denotes the service area of the broadcast by the data superimposing broadcast apparatus 3. Reference symbols A1, A2, A3 are the service areas of the broadcast by respective base stations of the interactive radio network 2. That is, a plurality of base stations exist in the service area Z of the broadcast by one data superimposing broadcast apparatus 3. The interactive radio network 2 is formed of a plurality of base stations. A plurality of base stations of the interactive radio network 2 may be overlapped on the service area by the different data superimposing broadcast apparatus 3. Furthermore, reference numbers 110, 111, 112, 113, 114 denote boundaries of the directivity of each directional antenna of the data superimposing broadcast apparatus 3, respectively.

The directivity of the individual user accessing the Web server 6 on internet 5 in the service area A1 may be overlapped on that of the individual user accessing in the

service area A2. Therefore, although it is difficult to ensure a sufficient CIR, there is a possibility that the directivity of the user in the service area A1 may be separated from that of the user in the service area A3 by a sufficient directivity. In this case, the corresponding different data can be transmitted from the antennas corresponding to each service area to the individual user in the service area A1 and the individual user in the service area A3, respectively, with the same frequency at the same time. The individual user in the service area A1 and the individual user in the service area A3 can simultaneously receive the corresponding data.

When the data is transmitted to the individual user in the service area A1 alone, the directional antenna limited by the directional boundaries 110 and 111 is used, whereby it is not necessary to transmit the radio wave in any other unnecessary directions, so that a radio wave resource can be effectively used. Furthermore, since an unnecessary radio wave is not outputted, interference between the radio waves can be reduced.

According to the above embodiments, although only the browser data of internet 5 is described above, the present invention is not limited to this. The value-added network 8 is the same as above. For example, the present invention can be applied to a television learning system, news (selection of detailed items), advertisement (requirement of contents explanation), a quiz/game (answer), a questionnaire and the like. Furthermore, if the encryption technique and the individual ID are used, an education system, a home shopping and the like to the specific individual are possible.

As described above, according to the present invention, the following effects can be obtained.

(1) The information terminal can perform the interactive communication. When respective information amounts to be transmitted in up link and down link are asymmetric to each other, i.e., they are different from each other, more specifically, when the down link information amount is larger than the up link information amount, the large amount data is transmitted from means for transmitting the large amount data which can transmit the large amount data. Therefore, data transmission can be so effectively carried out as to make the most of a high speed of the means for transmitting the large amount data, whereby the load applied to the interactive radio network can be reduced.

(2) In the interactive radio network, the base station which interactively transmits/receives the data relative to the information terminal existing in the corresponding service area is located in each predetermined service area. Therefore, as the interactive radio network, for example, existing PHS and cellular telephone (portable telephone) can be used.

(3) Since the large amount data is usually superimposed to the idle area of the television signal to be broadcasted and it is broadcasted, a usual television broadcasting station can be used as the means for transmitting the large amount data.

(4) In the means for transmitting the large amount data, the large amount data to be transmitted is temporarily stored in the memory, whereby the scheduled time of broadcasting can be set according to the data amount of the large amount data. Therefore, a transmitting schedule can be determined. Furthermore, the scheduled time of broadcasting is supplied to the information terminal, whereby the information terminal may operate means for receiving the large amount data during the time zone when the large amount data is transmitted. Therefore, the power consumption can be reduced.

(5) Since appropriate means for transmitting the large amount data can be selected from a plurality of means based on the present position of the information terminal, the radio wave resource can be effectively used.

(6) A predetermined encryption is carried out relative to the data transmitted/received between the information terminal and the interactive radio network and the data transmitted from the mean for transmitting the large amount data to the information terminal. Therefore, the leakage of the information can be prevented. Furthermore, the specific information terminal alone which obtains the encrypting key can receive (decrypt) the data. Therefore, for example, a charged service necessary for the charge management can be provided.

(7) When the server on internet is assumed as information providing means, the asymmetric router comprises browsing means for browsing the information (the large amount data) on the server. The information terminal carried out only such processes as to transmit the command and to display the predetermined format display data. Therefore, the load applied to the memory amount and the processing rate of the information terminal can be reduced. Furthermore, a circuit scale can be reduced. The portability is not lost, and a comfortable usage environment can be provided.

(8) A radio network information indicative of the available interactive radio network in the corresponding service area is transmitted by the means for transmitting the large amount data. Therefore, even if the information terminal is moved to a different service area (a different interactive radio network or different means for transmitting the large amount data), an appropriate system can be selected, and the information terminal can be moved over the wide range.

(9) An identification code for specifying the information terminal is added to the large amount data transmitted by the means for transmitting the large amount data. Therefore, the information terminal can receive only the large amount data related to the information terminal. The leakage of the information can be prevented, and the charged service can be realized.

(10) The means for transmitting the large amount data comprises a plurality of directional antennas having the directivity in different directions. Furthermore, the information terminal supplies the position information indicative of the present position to the means for transmitting the large amount data via the interactive radio network and the asymmetric router at a predetermined timing. Therefore, a number of data can be simultaneously transmitted to a plurality of information terminals with the same frequency. Furthermore, the radio wave resource can be effectively used, and a transmitting efficiency can be improved.

Additional advantages and modifications will readily occur to those skilled in the art. Therefore, the present invention in its broader aspects is not limited to the specific details, representative devices, and illustrated examples shown and described herein. Therefore, various modifications may be made without departing from the spirit or scope of the general inventive concept as defined by the appended claims and their equivalents.

We claim:

1. A data communication system comprising:

an interactive communication system in which data is transmitted/received between a host server and a terminal;

a large amount data transmitting system for unidirectionally transmitting a large amount of data from the host server to said terminal; and

a data supplying apparatus for supplying data to be transmitted to said terminal to said interactive communication system when an amount of the data is less than a predetermined amount, and for supplying the data to be transmitted to said terminal to said large amount data

19

transmitting system when the amount of the data is larger than the preset amount;

wherein said large amount data transmitting system comprises:

- (i) storage means for temporarily storing large amount data supplied by the data supplying apparatus;
- (ii) a timer for determining a scheduled transmission timing; and
- (iii) transmitting means for transmitting the large amount data stored in said storage means to the terminal at the scheduled transmission timing.

2. The communication system according to claim 1, wherein said transmitting means of said large amount data transmitting system comprises a broadcasting system for broadcasting a television signal, and means for superimposing the large amount data to be transmitted to said terminal on said television signal so that the large amount data is broadcasted to the terminal.

3. The communication system according to claim 1, wherein:

said large amount data transmitting system further comprises means for supplying data indicating the scheduled transmission timing to said terminal via said interactive communication system; and

said terminal comprises means for receiving the data indicating the scheduled transmission timing, and means for turning on a receiving circuit at the scheduled transmission timing so that the terminal may receive the large amount data transmitted by said large amount data transmitting system.

4. The communication system according to claim 1, wherein:

said large amount data transmitting system comprises a plurality of data transmitting devices which are located in respective predetermined areas; and

said data supplying apparatus comprises means for detecting a position of the terminal and means for supplying the data to be transmitted to the terminal to a data transmitting device located in an area corresponding to the position of said terminal.

5. The communication system according to claim 1, wherein:

said data supplying apparatus performs a predetermined encryption using a predetermined encrypting key relative to the large amount data supplied to said large amount data transmitting system; and

said terminal includes means for decrypting encrypted large amount data transmitted by said large amount data transmitting system using the predetermined encrypting key.

6. The communication system according to claim 5, wherein said terminal obtains the predetermined encryption key via said interactive communication system prior to a reception of the encrypted large amount data.

7. The communication system according to claim 1, wherein said interactive communication system and said terminal perform a predetermined encryption using a predetermined encrypting key relative to small amount data to

20

be transmitted/received, and the encrypted small amount data is decrypted using the predetermined encrypting key.

8. The communication system according to claim 1, wherein:

said host server comprises a plurality of internet servers for providing an HTML formatted file; and

said data supplying apparatus browses the HTML formatted file provided by a predetermined internet server according to a command supplied from said terminal via said interactive communication system, and converts the HTML formatted file into display format data which is transmitted to said large amount data transmitting system.

9. The communication system according to claim 4, wherein:

said large amount data transmitting system transmits communication information indicating a type of interactive communication device which can be used in each area by said terminal; and

said terminal selects a specific interactive communication device to be used according to the communication information transmitted by said large amount data transmitting system.

10. The communication system according to claim 1, wherein:

said large amount data transmitting system transmits an identification code for specifying the terminal to receive the large amount data; and

said terminal determines whether or not the identification code provided by said large amount data transmitting system corresponds to a preset identification code of the terminal, and enables the large amount data to be received only if a correspondence is detected.

11. The communication system according to claim 1, wherein:

said terminal supplies position information indicative of a present position of the terminal to said interactive communication system at a predetermined timing;

said data supplying apparatus transmits the position information via said interactive communication system to said large amount data transmitting system; and

said large amount data transmitting system comprises a plurality of transmitting antennas each having a different directivity, so that the large amount data may be selectively transmitted from a given one of the data transmission antennas having a directivity corresponding to said position information.

12. The communication system according to claim 1, wherein said interactive communication system comprises a base station connected to the terminal via a radio channel and a wired network connected between the base station and the host server.

13. The communication system according to claim 1, wherein said interactive communication system comprises means for transmitting a data request from the terminal to the host server.

* * * * *